

# CENTRAL LIMIT THEOREM FOR BIASED RANDOM WALK ON MULTI-TYPE GALTON–WATSON TREES

\*AMIR DEMBO AND †NIKE SUN

**ABSTRACT.** Let  $\mathcal{T}$  be a rooted supercritical multi-type Galton–Watson (MGW) tree with types coming from a finite alphabet, conditioned to non-extinction. The  $\lambda$ -biased random walk  $(X_t)_{t \geq 0}$  on  $\mathcal{T}$  is the nearest-neighbor random walk which, when at a vertex  $v$  with  $d_v$  offspring, moves closer to the root with probability  $\lambda/(\lambda + d_v)$ , and to each of the offspring with probability  $1/(\lambda + d_v)$ . This walk is recurrent for  $\lambda \geq \rho$  and transient for  $0 \leq \lambda < \rho$ , with  $\rho$  the Perron–Frobenius eigenvalue for the (assumed) irreducible matrix of expected offspring numbers. Subject to finite moments of order  $p > 4$  for the offspring distributions, we prove the following quenched CLT for  $\lambda$ -biased random walk at the critical value  $\lambda = \rho$ : for almost every  $\mathcal{T}$ , the process  $|X_{[nt]}|/\sqrt{n}$  converges in law as  $n \rightarrow \infty$  to a reflected Brownian motion rescaled by an explicit constant. This result was proved under some stronger assumptions by Peres–Zeitouni (2008) for single-type Galton–Watson trees. Following their approach, our proof is based on a new explicit description of a reversing measure for the walk from the point of view of the particle (generalizing the measure constructed in the single-type setting by Peres–Zeitouni), and the construction of appropriate harmonic coordinates. In carrying out this program we prove moment and conductance estimates for MGW trees, which may be of independent interest. In addition, we extend our construction of the reversing measure to a biased random walk with random environment (RWRE) on MGW trees, again at a critical value of the bias. We compare this result against a transience–recurrence criterion for the RWRE generalizing a result of Faraud (2011) for Galton–Watson trees.

## 1. INTRODUCTION

Let  $\mathcal{T}$  denote an infinite tree with root  $o$ . The  $\lambda$ -biased random walk on  $\mathcal{T}$ , hereafter denoted  $\text{RW}_\lambda(\mathcal{T})$ , is the Markov chain  $(X_t)_{t \geq 0}$  with  $X_0 = o$  such that given  $X_t = v$  with offspring number  $d_v$  and  $v \neq o$ ,  $X_{t+1}$  equals the parent of  $v$  with probability  $\lambda/(\lambda + d_v)$ , and is uniformly distributed among the offspring of  $v$  otherwise (and if  $X_t = o$ , then  $X_{t+1}$  is uniformly distributed among the offspring of  $o$ ).

For supercritical Galton–Watson trees without leaves, if  $\rho$  denotes the mean offspring number, then  $\text{RW}_\lambda$  is a.s. recurrent if and only if  $\lambda \geq \rho$  ([24, Thm. 4.3 and Propn. 6.4]), and ergodic if and only if  $\lambda > \rho$  ([19, Propn. 9-131] and [24, p. 944 and

---

*Date:* May 8, 2012.

*2010 Mathematics Subject Classification.* Primary 60F05, 60K37; Secondary 60J80, 60G50.

*Key words and phrases.* Multi-type Galton–Watson tree, biased random walk, central limit theorem, random walk with random environment.

\*†Research partially supported by NSF grants DMS-0806211 and DMS-1106627. †Research partially supported by NSF grant DMS-0502385 and Department of Defense NDSEG Fellowship.

p. 954]). With  $|v|$  denoting the (graph) distance from vertex  $v$  to the root  $o$ ,  $|X_t|/t$  converges a.s. to a speed  $V$ , with  $V = V(\lambda)$  deterministic, positive for  $\lambda < \rho$  and zero otherwise (see [26, 27] for  $\lambda < \rho$  and [31] for  $\lambda = \rho$ ; the case  $\lambda > \rho$  follows trivially from positive recurrence).

Further, subject to no leaves and finite exponential moments for the offspring distribution, a quenched CLT for  $\text{RW}_\lambda$  ( $\lambda \leq \rho$ ) on single-type Galton–Watson trees was shown by Peres–Zeitouni [31], and extended to the setting of random walk with random environment (RWRE) by Faraud [12]. In contrast, if leaves occur, there emerges a zero-speed transient regime  $\lambda < \lambda_c$  (for  $\lambda_c < \rho$ ) [27] where the leaves “trap” the random walk and create slow-down. It follows from the results of Ben Arous et al. [2] that in this setting, for sufficiently small  $\lambda$  there cannot be a (functional) CLT with diffusive scaling. Analogous results on the critical ( $\rho = 1$ ) Galton–Watson tree conditioned to survive were shown by Croydon et al. [8]. In this paper we consider the critical case  $\lambda = \rho$ , where [31, Thm. 1] proves that on a.e. Galton–Watson tree, the processes  $(|X_{\lfloor nt \rfloor}|/\sqrt{n})_{t \geq 0}$  converge in law to the absolute value of a (deterministically) scaled Brownian motion. Their proof is based on the construction of harmonic coordinates and an explicit description of a reversing probability measure IGWR for  $\text{RW}_\rho$  “from the point of view of the particle.” Having such an explicit description is a very delicate property: even for Galton–Watson trees, no such description is known for  $\lambda < \rho$  except at  $\lambda = 1$  which is done by [26, Thm. 3.1]. One thus might be led to believe that [31, Thm. 1] is a particular property resulting from the independence inherent in the Galton–Watson law.

Here we show to the contrary that such a quenched CLT extends to the much larger family of supercritical multi-type Galton–Watson trees with finite type space. We allow for leaves (but condition on non-extinction), demonstrating that at  $\lambda = \rho$  the “trapping” phenomenon of [2] does not arise. We also replace the assumption of exponential moments for the offspring distribution by an assumption of finite moments of order  $p > 4$ , so that our result restricted to the single-type case strengthens [31, Thm. 1]. However, the main interest of our result lies in moving from an i.i.d. to a Markovian structure for the random tree.

As in [31], the key ingredient in our proof is the construction of an explicit reversing (probability) measure IMGWR for  $\text{RW}_\lambda$  from the point of view of the particle, generalizing IGWR to the multi-type setting, for  $\lambda$  at the critical value on the boundary between transience and recurrence. See §2 for the details of the construction which may be of independent interest.

The model we consider is as follows: let  $\Omega$  be the space of rooted trees with type, where each vertex  $v$  is given a type  $\chi_v$  from a finite alphabet  $\mathcal{Q}$ . We let  $\mathcal{B}_\Omega$  be the  $\sigma$ -algebra on  $\Omega$  generated by the cylinder sets (determined by the restrictions of trees to finite neighborhoods of the root). We write  $\mathcal{T}$  for a generic element of  $\Omega$  and  $o$  for its root. A *multi-type Galton–Watson* tree is a random element  $\mathcal{T} \in \Omega$ , generated from a starting type  $\chi_o \in \mathcal{Q}$  and a collection of probability measures  $\mathbf{q}^a$  ( $a \in \mathcal{Q}$ ) on

$$\mathcal{Q}^* \equiv \bigcup_{\ell \geq 0} \mathcal{Q}^\ell,$$

as follows: begin with a root vertex  $o$  of type  $\chi_o$ . Supposing inductively that the first  $n$  levels of  $\mathcal{T}$  have been constructed, each vertex  $v$  at the  $n$ -th level generates random offspring according to law  $\mathbf{q}^{x_v}$ . For our purposes the ordering of the children does not matter, so each  $\mathbf{q}^a$  may equivalently be regarded as a probability measure on configurations  $\underline{x} = (x_b)_{b \in \mathcal{Q}} \in (\mathbb{Z}_{\geq 0})^{\mathcal{Q}}$ , where  $x_b$  is the number of children of type  $b$ . Continuing to construct successive generations in this Markovian fashion, we denote the resulting law on  $(\Omega, \mathcal{B}_{\Omega})$  by  $\text{MGW}^{\chi_o}$ . We denote by  $\text{MGW}$  any mixture of the measures  $(\text{MGW}^a)_{a \in \mathcal{Q}}$  (with (2.1) the canonical mixture) and let  $\mathbb{X} \equiv \{|\mathcal{T}| < \infty\}$  denote the event of extinction.

For  $a, b \in \mathcal{Q}$  let

$$A(a, b) = \sum_{\underline{x}} \mathbf{q}^a(\underline{x}) x_b,$$

the expected number of offspring of type  $b$  at a vertex of type  $a$ . (Unless otherwise specified, the implicit assumption hereafter is that  $\mathbb{E}_{\mathbf{q}^a}[|\underline{x}|] < \infty$  for all  $a \in \mathcal{Q}$  where  $|\underline{x}| \equiv \sum_b x_b$ .) Throughout the paper we will refer to the following assumptions:

(H1) The matrix  $A \equiv (A(a, b))_{a, b \in \mathcal{Q}}$  is irreducible with Perron–Frobenius eigenvalue  $\rho$ .

(H2)  $A$  is positive regular (every entry of  $A^{n_0}$  is positive for some  $n_0 \in \mathbb{N}$ ),  $\rho > 1$ , and  $\mathbb{E}_{\mathbf{q}^a}[|\underline{x}| \log |\underline{x}|] < \infty$  for all  $a \in \mathcal{Q}$ .

(H3<sup>p</sup>)  $\mathbb{E}_{\mathbf{q}^a}[|\underline{x}|^p] < \infty$  for all  $a \in \mathcal{Q}$ .

Note that (H1) and  $\rho > 1$  together imply  $\text{MGW}^a(\mathbb{X}) < 1$  for all  $a \in \mathcal{Q}$ .

**1.1. Central limit theorems.** We take all real-valued processes to be in the space  $D[0, \infty)$  equipped with the topology of uniform convergence on compact intervals. Our main theorem is the following:

**Theorem 1.1.** *Under (H1), (H2), and (H3<sup>p</sup>) with  $p > 4$ , for MGW-a.e.  $\mathcal{T} \notin \mathbb{X}$ , if  $X \sim \text{RW}_{\rho}(\mathcal{T})$  then the processes  $(|X_{\lfloor nt \rfloor}|/(\sigma\sqrt{n}))_{t \geq 0}$  converge in law in  $D[0, \infty)$  to the absolute value of a standard Brownian motion for  $\sigma$  a deterministic positive constant (see (3.1)).*

**Remark 1.2.** By [11, Propn. 3.10.4], an equivalent statement is that the polygonal interpolation of  $k/n \mapsto |X_k|/(\sigma\sqrt{n})$  converges to standard Brownian motion in the space  $C[0, \infty)$  (again with the topology of local uniform convergence).

Let  $\text{RW}_{\lambda}^{\text{cts}}(\mathcal{T})$  denote the continuous-time version of  $\text{RW}_{\lambda}(\mathcal{T})$ , which when at  $v \in \mathcal{T}$  moves to the parent of  $v$  (if  $v \neq o$ ) at rate  $\lambda$  and to each offspring of  $v$  at rate 1.

**Corollary 1.3.** *Under the assumptions of Thm. 1.1, for MGW-a.e.  $\mathcal{T} \notin \mathbb{X}$ , if  $X^{\text{cts}} \sim \text{RW}_{\rho}^{\text{cts}}(\mathcal{T})$  then the processes  $(|X_{nt}^{\text{cts}}|/(\sigma\sqrt{2\rho n}))_{t \geq 0}$  converge in law in  $D[0, \infty)$  to the absolute value of a standard Brownian motion.*

By moving the root of the tree to the current position of the random walk,  $\text{RW}_{\lambda}$  on the tree induces a random walk on the space  $\Omega$ , the “walk from the point of view of the particle.” As in [31, §3], to make the latter process Markovian we amend the state space so as to keep track of the ancestry of the vertices. Specifically, we consider

the space  $\Omega_\downarrow$  of pairs  $(\mathcal{T}, \xi)$ , where  $\mathcal{T}$  is an infinite tree and  $\xi = (o = v_0, v_1, v_2, \dots)$  is a ray emanating from the root  $o$ ; this ray indicates the ancestry of each vertex in the tree. Let  $\mathcal{B}_{\Omega_\downarrow}$  denote the  $\sigma$ -algebra generated by the cylinder sets. We define a height function  $h$  on  $\mathcal{T}$  as follows: set  $h(v_n) = -n$ , and for  $v \notin \xi$  set

$$h(v) = h(R_v) + d(v, \xi) \quad (1.1)$$

where  $d$  denotes graph distance and  $R_v$  is the nearest vertex to  $v$  on  $\xi$  (see Fig. 1). We denote by  $\text{RW}_\lambda(\mathcal{T}, \xi)$  the  $\lambda$ -biased random walk  $(Y_t)_{t \geq 0}$  on  $(\mathcal{T}, \xi)$ , where the bias goes in the direction of decreasing height. With  $\mathcal{T}^v$  the tree  $\mathcal{T}$  rooted at  $v$  instead of  $o$ , and  $\xi^v$  the unique ray emanating from  $v$  such that  $\xi \cap \xi^v$  is an infinite ray, let

$$(\mathcal{T}, \xi)^{Y_t} \equiv (\mathcal{T}^{Y_t}, \xi^{Y_t}), \quad t \geq 0. \quad (1.2)$$

This is a Markov process with state space  $\Omega_\downarrow$ , and we hereafter refer to it as  $\text{TRW}_\lambda$ . Let  $\text{RW}_\lambda^{\text{cts}}$  denote the continuous-time version of  $\text{RW}_\lambda(\mathcal{T}, \xi)$  (moving in the direction of increasing height at rate 1 and in the direction of decreasing height at rate  $\lambda$ ), and let  $\text{TRW}_\lambda^{\text{cts}}$  denote the induced continuous-time process on the space  $\Omega_\downarrow$ .

As in the single-type Galton–Watson case considered in [31], the key to our proof lies in finding an explicit reversing measure  $\text{IMGW}$  for  $\text{TRW}_\rho^{\text{cts}}$ , which is then easily translated to a reversing measure  $\text{IMGWR}$  for  $\text{TRW}_\rho$ . For a tree  $\mathcal{T}$  (with or without marked ray) and for any vertex  $v \in \mathcal{T}$ , we denote by  $\mathcal{T}^{(v)}$  the subtree induced by  $v$  and its descendants, where descent is in direction of increasing distance from the root for a rooted tree, and in the direction of increasing height for a tree with marked ray. If  $\mu$  is a law on trees we use  $\mu \otimes \text{RW}_\lambda$  to denote the joint law of the tree together with the realization of  $\text{RW}_\lambda$  on that tree.

**Theorem 1.4.** *Assume (H1).*

(a) *There exists a reversing probability measure  $\text{IMGW}$  for  $\text{TRW}_\rho^{\text{cts}}$ , and if we define*

$$\frac{d\text{IMGWR}}{d\text{IMGW}} = \frac{d_o + \rho}{2\rho},$$

*then  $\text{IMGWR}$  is a reversing probability measure for  $\text{TRW}_\rho$ .*

(b) *For  $((\mathcal{T}, \xi), (Y_t)_{t \geq 0}) \sim \text{IMGWR} \otimes \text{RW}_\rho$ , the stationary sequence  $((\mathcal{T}, \xi)^{Y_t})_{t \geq 0}$  is ergodic.*

The  $\text{IMGW}$  trees always have an infinite ray  $\xi$ , though the trees coming off the ray may be finite. The measures  $\text{IMGW}, \text{IMGWR}$  are the multi-type analogues of the measures  $\text{IGW}, \text{IGWR}$  of [31]. Thm. 1.4 and the construction of harmonic coordinates allow us to prove the following quenched CLT for  $\text{RW}_\rho$  on  $\text{IMGWR}$  trees, which will be used to deduce Thm. 1.1.

**Theorem 1.5.** *Under (H1), (H2), and (H3<sup>p</sup>) with  $p > 2$ , for  $\text{IMGWR}$ -a.e.  $(\mathcal{T}, \xi)$ , if  $Y \sim \text{RW}_\rho(\mathcal{T}, \xi)$  then the processes  $(h(Y_{[nt]})/(\sigma\sqrt{n}))_{t \geq 0}$  converge in law in  $D[0, \infty)$  to a standard Brownian motion.*

**1.2. Transience–recurrence boundary in random environment.** In the setting of  $\text{RW}_\lambda$  on  $\text{MGW}$  trees,  $\lambda = \rho$  represents the onset of recurrence. Indeed,  $\text{MGW}$ -a.e. tree  $\mathcal{T}$  on the event of non-extinction has branching number  $\text{br } \mathcal{T} = \rho$  [24, Propn. 6.5], therefore  $\text{RW}_\lambda(\mathcal{T})$  is transient for  $\lambda < \rho$  and recurrent for  $\lambda > \rho$  [24, Thm. 4.3]. In fact, recurrence for all  $\lambda \geq \rho$  follows from a simple conductance calculation (for the general theory see [28, Ch. 2]), therefore  $\rho$  is the boundary between transience and recurrence for  $\text{RW}_\lambda$  on  $\text{MGW}$  trees. Further  $\rho$  is the boundary between non-ergodicity and ergodicity, with  $\text{RW}_\rho$  null recurrent (see [19, Propn. 9-131] and [24, p. 944 and p. 954]) and of zero speed (e.g. from the bound of Lem. 3.5).

We believe that the existence of a reversing measure and CLT is a feature of the onset of recurrence in a more general setting. Indeed, suppose each vertex  $v \in \mathcal{T} \setminus \{o\}$  has, in addition to its type  $\chi_v$  from the (finite) alphabet  $\mathcal{Q}$ , a weight  $\alpha_v \in (0, \infty)$ . Fixing such a tree  $\mathcal{T}$  (the environment), the  $\lambda$ -biased random walk with random environment  $\text{RWRE}_\lambda(\mathcal{T})$  is the Markov chain  $(X_t)_{t \geq 0}$  with  $X_0 = o$  which, when at vertex  $v$  with offspring

$$(\underline{y}, \underline{\alpha}) \equiv ((y_1, \alpha_1), \dots, (y_\ell, \alpha_\ell)) \in \bar{\mathcal{Q}}^\ell,$$

jumps to a random neighbor  $w$  of  $v$  with probability proportional to  $\alpha_w$  if  $w$  is a child of  $v$ , and to  $\lambda$  if  $w$  is the parent of  $v$ . (Note that  $\text{RW}_\rho(\mathcal{T})$  corresponds to the case  $\alpha_v = 1$  for all  $v$ .) We let  $\text{RWRE}_\lambda^{\text{cts}}(\mathcal{T})$  denote the continuous-time version of  $\text{RWRE}_\lambda(\mathcal{T})$ .

If  $\mathbf{q}^a$  ( $a \in \mathcal{Q}$ ) is a probability measure on

$$\bar{\mathcal{Q}}^\star \equiv \bigcup_{\ell \geq 0} \bar{\mathcal{Q}}^\ell, \quad \bar{\mathcal{Q}} = \mathcal{Q} \times (0, \infty),$$

then the collection  $(\mathbf{q}^a)_{a \in \mathcal{Q}}$  together with starting type  $\chi_o \in \mathcal{Q}$  specifies a law  $\overline{\text{MGW}}^{a_0}$  on the space  $\Omega$  of typed weighted rooted trees. As before we let  $\overline{\text{MGW}}$  denote any mixture of the  $\overline{\text{MGW}}^{a_0}$ . This model, studied in the single-type case in [12], allows for quite general distributions on the (immediate) neighborhood of each vertex, but conditioned on types the weights in different neighborhoods must be independent.

For  $\gamma \in \mathbb{R}$  and  $a, b \in \mathcal{Q}$ , let

$$\bar{A}^{(\gamma)}(a, b) \equiv \int_{\bar{\mathcal{Q}}^\star} \sum_j \mathbf{1}_{\{y_j=b\}} \alpha_j^\gamma d\mathbf{q}^a(\underline{y}, \underline{\alpha}) \quad (1.3)$$

(not necessarily finite for all  $\gamma$ ). Let  $\bar{\rho}(\gamma)$  be the Perron–Frobenius eigenvalue of  $\bar{A}^{(\gamma)}$  where well-defined (i.e. where  $\bar{A}^{(\gamma)}$  has finite entries and is irreducible), and  $\infty$  otherwise. We will prove the following characterization of the transience–recurrence boundary for  $\text{RWRE}_\lambda$ , extending part of [12, Thm. 1.1]:

**Theorem 1.6.** *Suppose  $\bar{A}^{(0)}$  is positive regular, and  $\bar{\rho}(\gamma) < \infty$  for  $\gamma$  in an open neighborhood of 0. For  $\lambda > 0$  let*

$$p_\lambda \equiv \inf_{0 \leq \gamma \leq 1} \frac{\bar{\rho}(\gamma)}{\lambda^\gamma}.$$

(a) *If  $p_\lambda < 1$ , then  $\text{RWRE}_\lambda$  is positive recurrent  $\overline{\text{MGW}}$ -a.s.*

(b) If  $p_\lambda > 1$ , then  $\text{RWRE}_\lambda$  is transient  $\overline{\text{MGW}}(\cdot | \mathbb{X}^c)$ -a.s.

Thus the transience–recurrence boundary for  $\text{RWRE}_\lambda$  occurs at the unique value  $\lambda = \rho^\circ$  for which  $p_{\rho^\circ} = 1$ . On the other hand, let  $\Omega_\downarrow$  denote the space of typed weighted trees with ray, and let  $\text{TRWRE}_\lambda$  and  $\text{TRWRE}_\lambda^{\text{cts}}$  denote the Markov chains in  $\Omega_\downarrow$  induced by  $\text{RWRE}_\lambda$  and  $\text{RWRE}_\lambda^{\text{cts}}$  respectively. We have the following generalization of Thm. 1.4 (a):

**Theorem 1.7.** *Suppose  $\overline{\text{MGW}}$  is such that  $\bar{A} \equiv \bar{A}^{(1)}$  is irreducible with Perron–Frobenius eigenvalue  $\bar{\rho} \equiv \bar{\rho}(1)$ . Then there exists a reversing probability measure  $\overline{\text{IMGW}}$  on  $\Omega_\downarrow$  for  $\text{TRWRE}_{\bar{\rho}}^{\text{cts}}$ . If we let  $\alpha_{0j}$  denote the weight for the  $j$ -th child of the root  $o$ , and set*

$$\frac{d\overline{\text{IMGWR}}}{d\overline{\text{IMGW}}} = \frac{\bar{\rho} + \sum_{j=1}^{d_o} \alpha_{0j}}{2\bar{\rho}},$$

then  $\overline{\text{IMGWR}}$  is a reversing probability measure for  $\text{TRWRE}_{\bar{\rho}}$ .

We can see that  $\rho^\circ$  matches  $\bar{\rho}$  if and only if the function  $\gamma \mapsto \bar{\rho}(\gamma)/(\rho^\circ)^\gamma$  attains its infimum over  $0 \leq \gamma \leq 1$  at  $\gamma = 1$ . If this fails, Thm. 1.7 still gives a reversing measure at  $\bar{\rho}$ , but  $\bar{\rho} > \rho^\circ$  and the walk is already positive recurrent above  $\rho^\circ$ . However, at least in the single-type case, we have  $\rho^\circ = \bar{\rho}$  in all cases in which a CLT is possible: indeed, if

$$\kappa \equiv \inf \left\{ \gamma \geq 0 : \frac{\bar{\rho}(\gamma)}{(\rho^\circ)^\gamma} = 1 \right\},$$

by results of [16] a CLT cannot hold unless  $\kappa \geq 2$  (see [12, p. 3]). We expect  $\kappa \geq 2$  also to be a necessary condition in the multi-type case, and thus Thm. 1.6 and Thm. 1.7 support the claim that reversing measures occur at the boundary between transience and recurrence *in cases in which a CLT is possible*. However, even in the single-type case the random environment creates technical difficulties, and the RWRE-CLT of [12] requires some restriction on  $\kappa$ . While we expect that the methods of this paper and [12] can also be adapted to extend the RWRE-CLT to the multi-type setting under the same restrictions on  $\kappa$ , new ideas are required to achieve a CLT for the entire regime  $\kappa \geq 2$ .

### Outline of the paper.

- In §2 we construct the reversing measure  $\text{IMGWR}$  for  $\text{TRW}_\rho$  (in §2.1) and its generalization  $\overline{\text{IMGWR}}$  for  $\text{TRWRE}_{\bar{\rho}}$  (in §2.2); these constructions are based on ideas from [21]. In §2.3 we give an alternative characterization of  $\text{IMGWR}$  (extending a characterization of [31] to the multi-type setting) which we use to prove ergodicity of the stationary sequence  $((\mathcal{T}, \xi)^{Y_t})_{t \geq 0}$ .
- In §3 we prove the quenched  $\text{IMGWR}$ -CLT Thm. 1.5: in §3.1 we construct on  $\text{IMGWR}$ -a.e.  $(\mathcal{T}, \xi)$  a function  $v \mapsto S_v$  ( $v \in \mathcal{T}$ ) which is harmonic with respect to the transition probabilities of  $\text{RW}_\rho(\mathcal{T}, \xi)$ . By stationary and ergodicity of  $((\mathcal{T}, \xi)^{Y_t})_{t \geq 0}$  with respect to  $\text{IMGWR}$  we are able to control the quadratic variation of the martingale  $M_t \equiv S_{Y_t}$  to obtain an  $\text{IMGWR}$ -a.s. martingale CLT.



In §3.2 we adapt the methods of [31] and [12] to show that  $h(Y_t)$  is uniformly well approximated by  $M_t/\eta$  ( $\eta$  an explicit constant), proving Thm. 1.5.

- In §4 we prove the quenched MGW-CLT Thm. 1.1. In §4.1 we review (a slight modification of) a construction of [31] which gives a “shifted coupling” of  $(\mathcal{T}, (X_t)_{t \geq 0}) \sim \text{MGW} \otimes \text{RW}_\rho$  with  $((\mathcal{T}^\circ, \xi), (Y_t)_{t \geq 0}) \sim \text{IMGW}_0 \otimes \text{RW}_\rho$  such that fresh excursions of  $X$  are matched with fresh excursions of  $Y$  away from  $\xi$ . From this we obtain an *annealed* MGW-CLT (in §4.2) for  $X$  by controlling the amount of time spent outside the coupled excursions as well as the drift of  $Y$  along  $\xi$ . Because of the dependence between  $\mathcal{T}$  and  $Y$  we do not see how to this coupling directly to prove a quenched (MGW-a.s.) CLT. Instead, in §4.3 we adapt the method of [5] to deduce Thm. 1.1 from the annealed CLT by controlling the correlation between two realizations of  $\text{RW}_\rho$  on a single MGW tree  $\mathcal{T}$  (as was done in [31, §7] in the case  $\lambda < \rho$ ).
- In §5 we prove Thm. 1.6 describing the transience–recurrence boundary for  $\text{RWRE}_\lambda$ . The main result needed is a large deviations estimate (Lem. 5.2) on the conductances at the  $n$ -th level of the tree.
- In §6 are collected some basic properties of MGW which are needed in the course of our proof and which may be of independent interest. In §6.2 we show that moments for the offspring distributions translate directly to moments for the normalized population size defined in §2.3. In §6.3 we prove the existence of harmonic moments for the normalized population size, and use this result to prove conductance estimates used in the proof of Thm. 1.1.

**Open problems.** We conclude this section by mentioning some open problems in this area. These problems are open even for single-type Galton–Watson trees.

- (1) Does a CLT with diffusive scaling hold for  $\text{RW}_\rho$  in the entire regime  $p \geq 2$ ?
- (2) Does a CLT with diffusive scaling hold for  $\text{RWRE}_{\bar{p}}$  in the entire regime  $\kappa \geq 2$ ?
- (3) What happens for simple random walk on the critical Galton–Watson tree (conditioned to survive)?
- (4) Does a CLT with any scaling (or other limit law) hold for  $\text{RW}_\rho$  when  $p < 2$ ?

A common feature of these problems is that while the reversing measure for the process from the perspective of the particle is given by Thm. 1.4, the method of martingale approximation used in [31, 12] and in this paper seem not to be directly applicable.

**Acknowledgements.** We are very grateful to Ofer Zeitouni for many helpful communications leading to the proof of ergodicity in Thm. 1.4 and the method of going from the annealed to quenched CLT in the proof of Thm. 1.1. A.D. thanks Alexander Fribergh for discussions on the results of [2] which motivated us to extend our results to trees with leaves. N.S. thanks Yuval Peres for several helpful conversations about the papers [31, 21] which led to the construction of the reversing measure. We thank the anonymous referee for many valuable comments on drafts of this paper.

## 2. REVERSING PROBABILITY MEASURES FOR $\text{TRW}_\rho$ AND $\text{TRWRE}_{\bar{\rho}}$

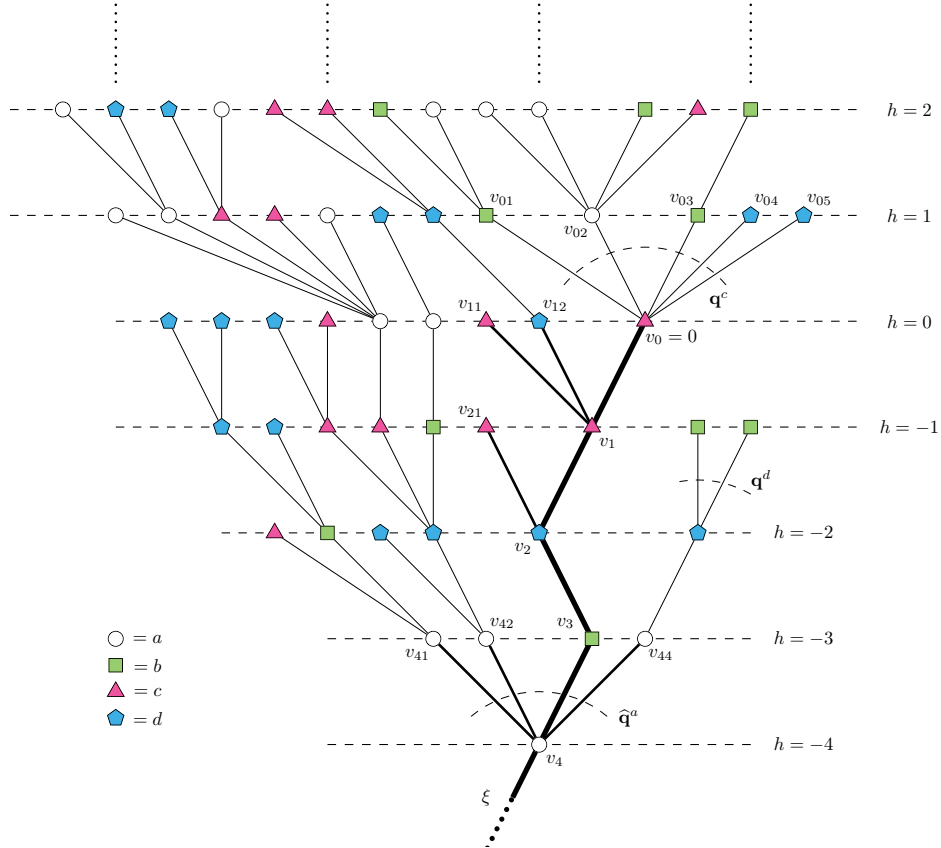
Assuming only (H1), in this section we construct the reversing measure  $\text{IMGWR}$  for  $\text{TRW}_\rho$  (§2.1) as well as its generalization  $\overline{\text{IMGWR}}$  for  $\text{TRWRE}_{\bar{\rho}}$  (in §2.2). In §2.3 we give an alternative characterization of  $\text{IMGWR}$  which we use to prove ergodicity of the stationary sequence  $((\mathcal{T}, \xi)^{Y_t})_{t \geq 0}$ . Except in §2.2 we work throughout with unweighted trees.

Consider a multi-type Galton–Watson measure  $\text{MGW}$  with offspring distributions  $(\mathbf{q}^a)_{a \in \mathcal{Q}}$  and mean matrix  $A$ . Hereafter we let  $\underline{e}$  and  $\underline{g}$  denote the right and left eigenvectors respectively associated to the Perron–Frobenius eigenvalue of  $A$ , normalized so that  $\sum_a g_a = \sum_a e_a = 1$ . Since our results are stated for  $\text{MGW}$ -a.e. tree, with no loss of generality we set hereafter

$$\mathbf{g}(a) \equiv \text{MGW}(\chi_o = a) = g_a. \quad (2.1)$$

Unless otherwise specified,  $X$  and  $Y$  denote  $\text{RW}_\rho$  on trees without and with marked ray respectively.

FIGURE 1.  $\text{IMGW}_0$  tree





**2.1. Construction of IMGWR.** We begin by constructing two auxiliary measures on the space  $\Omega_\downarrow$  of trees with ray  $(\mathcal{T}, \xi)$ . Let the infinite ray  $\xi$  (without types) be given. For some  $n > 0$ , we let vertex  $v_n$  be given a type  $\chi_n$  according to a distribution  $\pi$ , to be determined shortly. It is then given offspring  $\underline{x}^{v_n}$  according to the *inflated* offspring distribution  $\hat{\mathbf{q}}^{\chi_n}$ , where

$$\hat{\mathbf{q}}^a(\underline{x}) \equiv \mathbf{q}^a(\underline{x}) \frac{\langle \underline{x}, \underline{e} \rangle}{\rho e_a} \quad \forall a \in \mathcal{Q};$$

note that  $\hat{\mathbf{q}}^a(|\underline{x}| \geq 1) = 1$ . One offspring  $w$  of  $v_n$  is then identified with the next vertex  $v_{n-1}$  along  $\xi$ , where each  $w$  is chosen with probability  $e_{\chi_w} / \langle \underline{x}^{v_n}, \underline{e} \rangle$ . We proceed in this manner along the ray ending with the identification of  $v_0 = o$ . The sequence of types  $\chi_n, \chi_{n-1}, \dots$  seen along the ray is then (by (H1)) an irreducible Markov chain with transition probabilities

$$K(a, b) = \sum_{\underline{x}} \hat{\mathbf{q}}^a(\underline{x}) \frac{e_b x_b}{\langle \underline{x}, \underline{e} \rangle} = \sum_{\underline{x}} \mathbf{q}^a(\underline{x}) \frac{e_b x_b}{\rho e_a} = \frac{e_b}{\rho e_a} A(a, b). \quad (2.2)$$

This chain has stationary distribution  $\pi(a) \equiv e_a g_a / \langle \underline{e}, \underline{g} \rangle$ , so starting with  $\chi_n \sim \pi$  yields a consistent family of distributions for  $(v_n, \dots, v_1)$  and their (immediate) offspring, with types. By Kolmogorov's existence theorem, this uniquely specifies the distribution of the *backbone* of the tree, that is, of the ray  $\xi$  together with all (immediate) offspring of the vertices  $v_i$ ,  $i > 0$ . To each of these offspring (off the ray) and to  $o$ , we attach an independently chosen MGW tree conditioned on the given type, and denote by  $\text{IMGW}_0$  the resulting measure on  $\Omega_\downarrow$ .

The *inflated multi-type Galton–Watson* measure  $\text{IMGW}$  is obtained from  $\text{IMGW}_0$  by an additional biasing according to the root type  $\chi_o$ . Specifically, we set

$$\frac{d\text{IMGW}}{d\text{IMGW}_0} = \frac{1/e_{\chi_o}}{\mathbb{E}_\pi[1/e_\chi]} = \frac{\mathbb{E}_\mathbf{g}[e_\chi]}{e_{\chi_o}},$$

where  $\chi$  denotes a random variable on  $\mathcal{Q}$  with the specified distribution. We note that under  $\text{IMGW}$ ,  $\chi_o \sim \mathbf{g}$  and so  $\mathcal{T}^{(o)}$  has marginal law  $\text{MGW}$ , which implies

$$\mathbb{E}_{\text{IMGW}}[d_o] = \mathbb{E}_{\text{MGW}}[d_o] = \sum_a g_a \sum_b A(a, b) = \rho.$$

With this in mind, we define the probability measure  $\text{IMGWR}$  such that

$$\frac{d\text{IMGWR}}{d\text{IMGW}} = \frac{d_o + \rho}{\mathbb{E}_{\text{IMGW}}[d_o + \rho]} = \frac{d_o + \rho}{2\rho}, \quad (2.3)$$

and proceed to show that it is a reversing measure for  $\text{TRW}_\rho$ . From now on we adopt the notation that if  $\mu$  is a law on trees  $\mathcal{T}$  (with or without marked ray) and  $a \in \mathcal{Q}$ ,  $\mu^a$  refers to the law conditioned on  $\chi_o = a$ .

*Proof of Thm. 1.4 (a).* For the purposes of this proof we let  $\Omega$  and  $\Omega_\downarrow$  be spaces of *labelled* (or planar) trees (without and with marked ray, respectively), with corresponding Borel  $\sigma$ -algebras  $\mathcal{B}_\Omega$  and  $\mathcal{B}_{\Omega_\downarrow}$ . We extend  $\text{MGW}$ ,  $\text{IMGW}_0$ , etc. to be measures on these spaces by choosing an independent uniformly random ordering for the offspring of each vertex. For  $(\mathcal{T}, \xi) \in \Omega_\downarrow$  we use the shorthand  $i$  for  $v_i \in \xi$ , and write

$(i1, \dots, id_i)$  for its ordered offspring (with  $\underline{x}^i$  denoting the counts of offspring of  $i \equiv v_i$  of each type).

Recalling the notation of (1.2), let  $\mathcal{S}$  denote the map  $(\mathcal{T}, \xi) \mapsto (\mathcal{T}, \xi)^1$ . We will show that for  $A, B \in \mathcal{B}_{\Omega_\downarrow}$ ,

$$\int_A \mathbf{p}((\mathcal{T}, \xi), B) d\text{IMGWR}(\mathcal{T}, \xi) = \int_B \mathbf{p}((\mathcal{T}', \xi'), A) d\text{IMGWR}(\mathcal{T}', \xi') \quad (2.4)$$

where  $\mathbf{p}((\mathcal{T}, \xi), B)$  denotes the transition kernel of the process  $\text{TRW}_\rho$ . This identity implies reversibility of  $\text{TRW}_\rho$  on the space of labelled trees. Since this process projects to  $\text{TRW}_\rho$  on the space of unlabelled trees, the reversibility of the latter follows.

For  $(\mathcal{T}, \xi) \sim \text{IMGW}_0$ , let  $\text{IMGW}_0^a$  denote the law of the subtree  $\mathcal{T} \setminus \mathcal{T}^{(i-1)}$  rooted at  $i$  with marked ray  $\xi^i$ , conditioned on the event  $\{\chi_{i-1} = a\}$ , for any  $i \geq 1$  (note that this law does not depend on  $i$ ). Then

$$d\text{IMGW}_0(\mathcal{T}, \xi) = \pi(\chi_1) \frac{\hat{\mathbf{q}}^{\chi_1}(\underline{x}^1)}{d_1!} \frac{e_{\chi_0}}{\langle \underline{x}^1, \underline{e} \rangle} d\text{IMGW}_0^{\chi_1}(\mathcal{T} \setminus \mathcal{T}^{(1)}, \xi^2) \prod_{j=1}^{d_1} d\text{MGW}^{\chi_{1j}}(\mathcal{T}^{(1j)}).$$

Let  $\mathcal{P}^{\text{inj}}$  denote the collection of  $\mathcal{B}_{\Omega_\downarrow}$ -measurable sets on which  $\mathcal{S}$  is injective, and suppose  $B \in \mathcal{P}^{\text{inj}}$ . If  $\mu$  is a measure on  $\Omega_\downarrow$ ,  $\mathcal{S}_B^* \mu(\cdot) \equiv (\mu \circ \mathcal{S})(B \cap \cdot)$  is a well-defined measure on  $\Omega_\downarrow$ . Then

$$d\mathcal{S}_B^* \text{IMGW}_0(\mathcal{T}, \xi) = \mathbf{1}_{\{(\mathcal{T}, \xi) \in B\}} \pi(\chi_1) \frac{\mathbf{q}^{\chi_1}(\underline{x}^1)}{d_1!} d\text{IMGW}_0^{\chi_1}(\mathcal{T} \setminus \mathcal{T}^{(1)}, \xi^2) \prod_{j=1}^{d_1} d\text{MGW}^{\chi_{1j}}(\mathcal{T}^{(1j)}).$$

so

$$\frac{d\mathcal{S}_B^* \text{IMGW}_0}{d\text{IMGW}_0} = \mathbf{1}_B \frac{\rho e_{\chi_1}}{e_{\chi_0}}.$$

We then verify that

$$\begin{aligned} \frac{d\mathcal{S}_B^* \text{IMGW}}{d\text{IMGW}} &= \mathbf{1}_B \frac{\left( \frac{d\mathcal{S}_B^* \text{IMGW}}{d\mathcal{S}_B^* \text{IMGW}_0} \right)}{\left( \frac{d\text{IMGW}}{d\text{IMGW}_0} \right)} \frac{d\mathcal{S}_B^* \text{IMGW}_0}{d\text{IMGW}_0} \\ &= \mathbf{1}_B \frac{\left( \frac{d\text{IMGW}}{d\text{IMGW}_0} \circ \mathcal{S} \right)}{\left( \frac{d\text{IMGW}}{d\text{IMGW}_0} \right)} \frac{d\mathcal{S}_B^* \text{IMGW}_0}{d\text{IMGW}_0} = \mathbf{1}_B \frac{1/e_{\chi_1}}{1/e_{\chi_0}} \frac{\rho e_{\chi_1}}{e_{\chi_0}} = \mathbf{1}_B \rho, \end{aligned} \quad (2.5)$$

and similarly

$$\frac{d\mathcal{S}_B^* \text{IMGWR}}{d\text{IMGWR}} = \mathbf{1}_B \frac{\rho(d_1 + \rho)}{d_o + \rho}.$$

The left-hand side of (2.4) can be written as

$$\int_{A \cap \mathcal{S}^{-1}B} \frac{\rho}{d_o + \rho} d\text{IMGWR}(\mathcal{T}, \xi) + \int_A \frac{1}{d_o + \rho} \sum_{i=1}^{d_o} \mathbf{1}_{\{(\mathcal{T}^{0i}, \xi^{0i}) \in B\}} d\text{IMGWR}(\mathcal{T}, \xi).$$

Using the injectivity of  $\mathcal{S}$  on  $B$ , the second integral can be written as

$$\begin{aligned} \int_{A \cap \mathcal{S}B} \frac{1}{d_o + \rho} d\text{IMGWR}(\mathcal{T}, \xi) &= \int_{\mathcal{S}^{-1}A \cap B} \frac{1}{d_1 + \rho} d\mathcal{S}_B^* \text{IMGWR}(\mathcal{T}, \xi) \\ &= \int_{\mathcal{S}^{-1}A \cap B} \frac{\rho}{d_o + \rho} d\text{IMGWR}(\mathcal{T}, \xi). \end{aligned}$$

Combining these yields an expression for the left-hand side of (2.4) which is symmetric in  $A$  and  $B$ , from which it is clear that the two sides must agree.

Since every cylinder event  $F$  can be decomposed into the disjoint union of the event  $F_j = \{(\mathcal{T}, \xi) \in F : o = 1j\}$  (i.e.,  $o$  is the  $j$ -th child of 1), with  $F_j$  clearly in  $\mathcal{P}^{\text{inj}}$ , we have that  $\mathcal{P}^{\text{inj}}$  generates  $\mathcal{B}_{\Omega_\downarrow}$ . To conclude, for fixed  $A$  let  $\mathcal{B}'_{\Omega_\downarrow}$  denote the collection of sets  $B \in \mathcal{B}_{\Omega_\downarrow}$  for which (2.4) holds. From the above  $\mathcal{B}'_{\Omega_\downarrow}$  contains the  $\pi$ -system  $\mathcal{P}^{\text{inj}}$ . Further  $\mathcal{B}'_{\Omega_\downarrow}$  is closed under monotone limits and countable disjoint unions, and in particular it contains  $\Omega_\downarrow$  since  $\Omega_\downarrow$  can be decomposed as a countable disjoint union of sets in  $\mathcal{P}^{\text{inj}}$  by a similar argument as above. Thus by the  $\pi$ - $\lambda$  theorem (2.4) holds for all  $B \in \sigma(\mathcal{P}^{\text{inj}})$ , and extends to all  $B \in \mathcal{B}_{\Omega_\downarrow}$  again using the claim above.

The proof that  $\text{IMGW}$  is a reversing measure for the Markov pure jump process  $\text{TRW}_\rho^{\text{cts}}$  is similar: instead of (2.4) we show that

$$\int_A \lambda(\mathcal{T}, \xi) \mathbf{p}((\mathcal{T}, \xi), B) d\text{IMGW}(\mathcal{T}, \xi) = \int_B \lambda(\mathcal{T}, \xi) \mathbf{p}((\mathcal{T}', \xi'), A) d\text{IMGW}(\mathcal{T}', \xi') \quad (2.6)$$

where  $\lambda(\mathcal{T}, \xi) \equiv \lambda + d_o$  is the instantaneous jump rate of the process at state  $(\mathcal{T}, \xi)$ . As before, it suffices to show this for  $B \in \mathcal{P}^{\text{inj}}$ . In this case the left-hand side of (2.6) equals

$$\int_{A \cap \mathcal{S}^{-1}B} \rho d\text{IMGW}(\mathcal{T}, \xi) + \int_{\mathcal{S}^{-1}A \cap B} d\mathcal{S}_B^* \text{IMGW}(\mathcal{T}, \xi),$$

which by (2.5) coincides with the right-hand side of (2.6).  $\square$

**2.2. Extension of IMGWR to random environment.** We now extend the methods of the previous section to prove Thm. 1.7. Let  $\bar{e}, \underline{g}$  denote the right and left Perron–Frobenius eigenvectors of  $\bar{A} \equiv \bar{A}^{(1)}$ , normalized to have sum 1; as before we set  $\mathbf{g}(a) \equiv \bar{\text{MGW}}(\chi_o = a)$  to be  $\bar{g}_a$ .

We proceed much as in the deterministic environment setting, although the notation becomes more complicated. For  $\underline{y} \in \mathcal{Q}^\ell$  write  $\underline{\bar{e}}(\underline{y}) \equiv (\bar{e}_{y_j})_{j=1}^\ell$ . For  $a \in \mathcal{Q}$

$$\mathbb{E}_{\mathbf{q}^a}[\langle \underline{\bar{e}}(\underline{y}), \underline{\alpha} \rangle] = \sum_b \bar{A}(a, b) \bar{e}_b = \bar{\rho} \bar{e}_a,$$

so we define the inflated offspring measure  $\hat{\mathbf{q}}^a$  by

$$\frac{d\hat{\mathbf{q}}^a}{d\mathbf{q}^a} = \frac{\langle \underline{\bar{e}}(\underline{y}), \underline{\alpha} \rangle}{\bar{\rho} \bar{e}_a}.$$

We then construct the measure  $\bar{\text{IMGW}}_0$  on  $\Omega_\downarrow$  generalizing the measure  $\text{IMGW}_0$  of the previous section: let the infinite ray  $\xi$  (without types or weights) be given, and for some  $n > 0$  let  $v_n$  have type  $\chi_n$ . It is given offspring  $(\underline{y}^{v_n}, \underline{\alpha}^{v_n}) \sim \hat{\mathbf{q}}^{\chi_n}$ . One offspring

$w$  of  $v_n$  is identified with the next vertex  $v_{n-1}$  along  $\xi$ , where each  $w$  is chosen with probability

$$\frac{\bar{e}_{\chi_w} \alpha_w}{\langle \bar{e}(\underline{y}), \underline{\alpha}^{v_n} \rangle}.$$

Continuing the procedure along the ray up to  $v_0 = o$ , the sequence of types  $\chi_n, \chi_{n-1}, \dots$  seen along  $\xi$  is an irreducible Markov chain with transition probabilities

$$\bar{K}(a, b) = \mathbb{E}_{\bar{\mathbf{q}}^a} \left[ \frac{\bar{e}_b \sum_j \alpha_j \mathbf{1}_{\{y_j=b\}}}{\langle \bar{e}(\underline{y}), \underline{\alpha} \rangle} \right] = \frac{\bar{e}_b}{\bar{\rho} \bar{e}_a} \bar{A}(a, b)$$

and stationary distribution  $\bar{\pi}(a) = \bar{e}_a \bar{g}_a / \langle \bar{e}, \bar{g} \rangle$ . Thus, starting with  $\chi_n \sim \bar{\pi}$  and applying Kolmogorov's existence theorem, we obtain a measure  $\overline{\text{IMGW}}_0$  on  $\Omega_\downarrow$  which is a generalization of  $\text{IMGW}_0$ .

*Proof of Thm. 1.7.* The proof is by a straightforward modification of the proof of Thm. 1.4 (a). Let  $\mathcal{S} : (\mathcal{T}, \xi) \mapsto (\mathcal{T}, \xi)^1$ ; we emphasize that  $\mathcal{S}$  is a mapping on typed *weighted* labelled trees. For  $(\mathcal{T}, \xi) \sim \overline{\text{IMGW}}_0$ , let  $\overline{\text{IMGW}}_0^a$  denote the law of the subtree  $\mathcal{T} \setminus \mathcal{T}^{(i-1)}$  rooted at  $i$  with marked ray  $\xi^i$ , conditioned on the event  $\{\chi_{i-1}\}$ , for any  $i \geq 1$ . Let  $\mathcal{P}^{\text{inj}}$  denote the collection of  $\mathcal{B}_{\Omega_\downarrow}$ -measurable sets on which  $\mathcal{S}$  is injective. For  $A \in \mathcal{B}_{\Omega_\downarrow}$  and  $B \in \mathcal{P}^{\text{inj}}$ , we compute

$$\begin{aligned} d\overline{\text{IMGW}}_0(\mathcal{T}, \xi) &= \bar{\pi}(\chi_1) \hat{\mathbf{q}}^{\chi_1}(\underline{y}^1, \underline{\alpha}^1) \frac{\bar{e}_{\chi_0} \alpha_0}{\langle \bar{e}(\underline{y}^1), \underline{\alpha}^1 \rangle} d\overline{\text{IMGW}}_0^{\chi_1}(\mathcal{T} \setminus \mathcal{T}^{(1)}, \xi^2) \prod_{j=1}^{d_1} d\overline{\text{MGW}}^{\chi_{1j}}(\mathcal{T}^{(1j)}), \\ d\mathcal{S}_B^* \overline{\text{IMGW}}_0(\mathcal{T}, \xi) &= \mathbf{1}_{\{(\mathcal{T}, \xi) \in B\}} \bar{\pi}(\chi_1) \mathbf{q}^{\chi_1}(\underline{y}^1, \underline{\alpha}^1) d\overline{\text{IMGW}}_0^{\chi_1}(\mathcal{T} \setminus \mathcal{T}^{(1)}, \xi^2) \prod_{j=1}^{d_1} d\overline{\text{MGW}}^{\chi_{1j}}(\mathcal{T}^{(1j)}), \end{aligned}$$

so  $\bar{e}_{\chi_0} \alpha_0 d\mathcal{S}_B^* \overline{\text{IMGW}}_0 = \mathbf{1}_B \bar{\rho} \bar{e}_{\chi_1} d\overline{\text{IMGW}}_0$ . Letting

$$\begin{aligned} \frac{d\overline{\text{IMGW}}}{d\overline{\text{IMGW}}_0} &\equiv \frac{1/\bar{e}_{\chi_0}}{\mathbb{E}_{\bar{\pi}}[1/\bar{e}_\chi]} = \frac{\mathbb{E}_{\bar{\mathbf{g}}}[\chi]}{\bar{e}_{\chi_0}}, \\ \frac{d\overline{\text{IMGWR}}}{d\overline{\text{IMGW}}} &\equiv \frac{\bar{\rho} + \sum_{j=1}^{d_0} \alpha_{0j}}{\mathbb{E}_{\overline{\text{IMGW}}}[\bar{\rho} + \sum_{j=1}^{d_0} \alpha_{0j}]} = \frac{\bar{\rho} + \sum_{j=1}^{d_0} \alpha_{0j}}{2\bar{\rho}}, \end{aligned}$$

we obtain

$$\begin{aligned} \alpha_0 d\mathcal{S}_B^* \overline{\text{IMGW}} &= \mathbf{1}_B \bar{\rho} d\overline{\text{IMGW}}, \\ \frac{\alpha_0}{\bar{\rho} + \sum_{j=1}^{d_1} \alpha_{1j}} d\mathcal{S}_B^* \overline{\text{IMGWR}} &= \mathbf{1}_B \frac{\bar{\rho}}{\bar{\rho} + \sum_{j=1}^{d_0} \alpha_{0j}} d\overline{\text{IMGWR}}. \end{aligned}$$

The analogue of (2.4) thus holds for all  $B \in \mathcal{P}^{\text{inj}}$ , and we extend to all  $B \in \mathcal{B}_{\Omega_\downarrow}$  by essentially the same argument used in the proof of Thm. 1.4 (a).  $\square$

**2.3.  $\text{IMGW}_0$  as a weak limit and ergodicity.** In this section we provide an alternative characterization (Propn. 2.1) of the inflated Galton–Watson measure  $\text{IMGW}_0$ , which is then used in proving the ergodicity result Thm. 1.4 (b). Propn. 2.1 is also of independent interest as a multi-type extension of [31, Lem. 1].

To this end, we will define the notion of “normalized population size” for rooted trees  $\mathcal{T}$  with type. Let  $\mathcal{T}_n$  denote the subtree induced by  $\{v \in \mathcal{T} : |v| \leq n\}$ , and  $D_n$  the set  $\{v \in \mathcal{T} : |v| = n\}$ . Let  $(\mathcal{F}_n)_{n \geq 0}$  denote the natural filtration of the tree, i.e.,  $\mathcal{F}_n$  is the  $\sigma$ -algebra generated by  $\mathcal{T}_n$  (a finite tree with vertex types). Let  $\underline{Z}_n = (Z_n(b))_{b \in \mathcal{Q}} \in (\mathbb{Z}_{\geq 0})^{\mathcal{Q}}$  count the number of vertices of each type at level  $n$ , so  $\underline{Z}_n$  is  $\mathcal{F}_n$ -measurable. Then

$$\mathfrak{Z}_n \equiv \frac{\langle \underline{Z}_n, \underline{e} \rangle}{\rho^n} = \frac{1}{\rho^n} \sum_{v \in D_n} e_{\chi_v}$$

is a non-negative  $(\mathcal{F}_n)$ -martingale under  $\text{MGW}^a$  for every  $a$ , with  $\mathbb{E}_{\text{MGW}^a}[\mathfrak{Z}_0] = e_a$  (see e.g. [14, p. 49]). By the normalized population size of the tree we mean the a.s. limit of  $\mathfrak{Z}_n$ , denoted  $W_o$ . For  $v \in \mathcal{T}$  we use  $W_v$  to denote the normalized population size of  $\mathcal{T}^{(v)}$ . Under (H1) and (H2), it follows from the multi-type Kesten–Stigum theorem (see [20], or the conceptual proof of [21]) that  $W_o > 0$  a.s. on the event of non-extinction, and  $\mathbb{E}_{\text{MGW}^a}[W_o] = e_a$ .

For  $a \in \mathcal{Q}$  let  $\mathbb{Q}_n^a$  be a probability measure on (infinite) rooted trees defined by

$$\frac{d\mathbb{Q}_n^a}{d\text{MGW}^a} = \frac{\mathfrak{Z}_n}{e_a}. \quad (2.7)$$

For  $\mathcal{T} \sim \mathbb{Q}_n^a$  choose  $v_n \in D_n$  at random with probabilities proportional to weights  $e_{\chi_{v_n}}$ , and let  $\mathbb{Q}_{n*}^a$  denote the law of the resulting pair  $(\mathcal{T}, v_n)$ . Let  $\mathbb{Q}_{n*} \equiv \sum_{a \in \mathcal{Q}} \pi_a \mathbb{Q}_{n*}^a$  and  $\mathbb{Q}_n \equiv \sum_{a \in \mathcal{Q}} \pi_a \mathbb{Q}_n^a$ , so that  $d\mathbb{Q}_n/d\text{MGW} = \mathfrak{Z}_n/\mathbb{E}_{\mathbf{g}}[e_{\chi}]$ . Finally let  $\text{IMGW}_0(n)$  denote the law of  $(\mathcal{T}, \xi_0)^{v_n}$  (see (1.2) for this notation), where  $(\mathcal{T}, v_n) \sim \mathbb{Q}_{n*}$  and  $\xi_0$  is any infinite ray emanating from  $o$  not sharing an edge with the geodesic from  $o$  to  $v_n$ .

**Proposition 2.1.** *Under (H1),  $\text{IMGW}_0(n)$  converges weakly to  $\text{IMGW}_0$ .*

The proposition can be seen from the following explicit construction of  $\mathbb{Q}_{n*}^a$ : begin with  $v_0 \equiv o$  of type  $a$ , and suppose inductively that we have constructed  $(\mathcal{T}_i, v_i)$  ( $i < n$ ) where  $\mathcal{T}_i$  is the tree up to level  $i$  and  $v_i$  is the  $i$ -th vertex on the geodesic from  $o$  to  $v_n$ . Then  $v_i$  is given offspring  $\underline{x}^{v_i}$  according to  $\widehat{\mathbf{q}}^{x_{v_i}}$ , and one of these offspring  $w$  is randomly chosen (according to weights  $e_w$ ) to be distinguished as  $v_{i+1}$ . Meanwhile all other vertices  $v \in D_i \setminus \{v_i\}$  are given offspring  $\underline{x}^v$  according to  $\mathbf{q}^{x_v}$ . Once  $(\mathcal{T}_n, v_n)$  has been constructed, attach to each  $v \in D_n$  an independent  $\text{MGW}^{x_v}$  tree. For  $N \geq n$ ,

$$\frac{\mathbb{Q}_{n*}^a(\mathcal{T}_N, v_n)}{\text{MGW}^a(\mathcal{T}_N)} = \prod_{i=0}^{n-1} \frac{\langle \underline{x}^{v_i}, \underline{e} \rangle}{\rho e_{x_{v_i}}} \frac{e_{x_{v_{i+1}}}}{\langle \underline{x}^{v_i}, \underline{e} \rangle} = \frac{e_{x_{v_n}}}{\rho^n e_a},$$

and summing over  $v_n \in D_n$  gives (2.7).

Letting  $n \rightarrow \infty$  in  $\mathbb{Q}_{n*}^a, \mathbb{Q}_{n*}$  we obtain the measures  $\mathbb{Q}_{\infty*}^a, \mathbb{Q}_{\infty*}$  on rooted trees with infinite marked ray which coincide precisely with the measures  $\widehat{\text{MGW}}_{*}^a, \widehat{\text{MGW}}_{*}$  of [21].

The corresponding marginals  $Q_\infty^a \equiv \widehat{\text{MGW}}^a$ ,  $Q_\infty \equiv \widehat{\text{MGW}}$  on trees without marked ray satisfy

$$\left. \frac{dQ_\infty^a}{d\text{MGW}^a} \right|_{\mathcal{F}_n} = \frac{\mathfrak{Z}_n}{e_a}, \quad \left. \frac{dQ_\infty}{d\text{MGW}} \right|_{\mathcal{F}_n} = \frac{\mathfrak{Z}_n}{\mathbb{E}_{\mathbf{g}}[e_\chi]}.$$

By the Kesten–Stigum theorem and Scheffé’s lemma (see e.g. [34, §5.10]),  $\mathfrak{Z}_n \xrightarrow{L^1} W_o$ , hence

$$\frac{dQ_\infty^a}{d\text{MGW}^a} = \frac{W_o}{e_a}, \quad \frac{dQ_\infty}{d\text{MGW}} = \frac{W_o}{\mathbb{E}_{\text{MGW}}[W_o]} = \frac{W_o}{\mathbb{E}_{\mathbf{g}}[e_\chi]}.$$

We remark that although  $Q_{\infty\star} \equiv \widehat{\text{MGW}}_\star$  and  $\text{IMGWR}$  are both measures on trees with rays, they are not in general equivalent unless  $K$  is reversible.

*Proof of Propn. 2.1.* Since  $dQ_n/d\text{MGW} = \mathfrak{Z}_n/\mathbb{E}_{\mathbf{g}}[e_\chi]$  and  $\chi_o \sim \mathbf{g}$  under  $\text{MGW}$ , it follows that  $\chi_o \sim \pi$  under  $Q_{n\star}$ . It is then clear from the constructions of  $\text{IMGW}_0$  and  $Q_{n\star}$  that if  $(\mathcal{T}, \xi) \sim \text{IMGW}_0$ , then  $(\mathcal{T}^{(v_n)}, o) \sim Q_{n\star}$ . In other words the portion of  $(\mathcal{T}, \xi)$  descended from  $v_n$  has the same distribution under  $\text{IMGW}_0(n)$  as under  $\text{IMGW}_0$ , proving the result.  $\square$

Turning now to the proof of Thm. 1.4 (b), it is useful to define a two-sided version of  $\text{IMGW}_0$ , as follows. Let  $\Omega_\downarrow$  denote the space of trees with marked line: pairs  $(\mathcal{T}, \bar{\xi})$  where  $\mathcal{T}$  is an infinite tree and

$$\bar{\xi} \equiv (\dots, \bar{\xi}_{-1}, \bar{\xi}_0 = o, \bar{\xi}_1, \dots)$$

is a *line* (doubly infinite simple path) passing through the root. The positive and negative parts  $\bar{\xi}_\pm \equiv (\bar{\xi}_{\pm j})_{j \geq 0}$  of  $\bar{\xi}$  are edge-disjoint rays emanating from  $o$ .

Now suppose in the construction of  $\text{IMGW}_0$  we continue the backbone indefinitely rather than stopping at  $o$ , so that Kolmogorov’s existence theorem gives a doubly infinite backbone based on a line  $\bar{\xi}$ . Attaching  $\text{MGW}$  trees to the leaves of this backbone then gives a tree with marked line  $(\mathcal{T}, \bar{\xi})$ , whose law  $\underline{\text{IMGW}}$  is clearly stationary with respect to the shift  $\underline{\mathfrak{S}} : (\mathcal{T}, \bar{\xi}) \mapsto (\mathcal{T}, \bar{\xi})^{\xi_{-1}}$  which defined by moving the root to  $\bar{\xi}_{-1}$ . (Alternatively, if  $(\mathcal{T}, \xi)$  has law  $\text{IMGW}_0$  conditioned on non-extinction of  $\mathcal{T}^{(o)}$  and  $\bar{\xi}$  is any line with  $\bar{\xi}_- = \xi$ , then  $\underline{\mathfrak{S}}^n(\mathcal{T}, \bar{\xi})$  converges weakly to an  $\underline{\text{IMGW}}$  tree.)

It follows from the discussion preceding Propn. 2.1 that if we let

$$(\mathcal{T}, \xi) \sim \text{IMGW}_0 \quad \text{and} \quad (\mathcal{T}', \xi') \sim Q_{\infty\star}^{\chi_o}$$

(independently conditioned on  $\chi_o$ ), and we delete from  $\mathcal{T}$  all the vertices descended from  $o$  and identify  $o$  with the root of  $\mathcal{T}'$ , then we obtain a tree with marked line  $\bar{\xi}_- = \xi$ ,  $\bar{\xi}_+ = \xi'$  whose law is precisely  $\underline{\text{IMGW}}$ . It follows that the marginal law  $\text{IMGW}_1$  of  $(\mathcal{T}, \bar{\xi}_-)$  under  $\underline{\text{IMGW}}$  is given by

$$\frac{d\text{IMGW}_1}{d\text{IMGW}_0} = \frac{dQ_{\infty\star}^{\chi_o}}{d\text{MGW}^{\chi_o}} = \frac{W_o}{e_{\chi_o}}. \quad (2.8)$$

*Proof of Thm. 1.4 (b).* We adapt the proof of [35, Cor. 2.1.25]. Abbreviating  $\mathbb{T} \equiv (\mathcal{T}, \xi)$ , we let  $\nu$  denote the law of  $\underline{\mathbb{T}} \equiv (\mathbb{T}_t)_{t \geq 0} \equiv ((\mathcal{T}, \xi)^{Y_t})_{t \geq 0}$  in the space  $\Omega_\downarrow^\infty$  of

sequences of trees with ray, and  $\mathfrak{S}_0$  the shift  $(T_0, T_1, \dots) \mapsto (T_1, T_2, \dots)$  on  $\Omega_\downarrow^\infty$ . The content of the result is that the measure-preserving system  $(\Omega_\downarrow^\infty, \mathcal{F}^\infty, \nu, \mathfrak{S}_0)$  is ergodic.

*Step 1: reduction to induced system.*

Recall that under the measure  $\text{IMGW}_0$  the trees  $\mathcal{T}^{(i)} \setminus \mathcal{T}^{(i-1)}$  are conditionally independent given the ray  $\xi$  with types, and  $\max_{a \in \mathcal{Q}} \text{MGW}^a(\mathbb{X}) < 1$ ; therefore it holds  $\text{IMGW}_0$ -a.s. that  $|\mathcal{T}^{(i)}|$  for infinitely many  $i \in \xi$ . Since the walk  $Y_t$  on  $(\mathcal{T}, \xi)$  has a backward drift along  $\xi$ , this implies that if we let

$$A \equiv \{\underline{T} \in \Omega_\downarrow^\infty : T_0 = (\mathcal{T}, \xi) \text{ with } |\mathcal{T}^{(o)}| = \infty\}$$

and  $n_A(\underline{T}) \equiv \inf\{n \geq 1 : \mathfrak{S}_0^n \underline{T} \in A\}$  the first hitting time of  $A$  after time zero, then  $\nu(n_A < \infty) = 1$ . Thus  $(\Omega_\downarrow^\infty, \nu, \mathfrak{S}_0)$  forms a (Kakutani) tower over the induced measure-preserving system  $(A, \nu_A \equiv \nu(\cdot | A), \mathfrak{S}_0^{n_A})$ . We now show that the induced system is ergodic, which is equivalent to ergodicity of the original system ([32]; see also [26, §2]).

*Step 2: reduction to  $\mathcal{S}$ -invariance.*

Let  $n_A^i(\underline{T})$  denote the  $i$ -th hitting time of  $A$  after time zero and  $\mathcal{H}_i \equiv \sigma(T_0, \dots, T_{n_A^i})$ ; note that  $(T_{n_A^i})_{i \geq 0}$  forms an  $(\mathcal{H}_i)$ -Markov chain. Write  $\mathfrak{S} \equiv \mathfrak{S}_0^{n_A}$  and let  $\mathcal{I}$  denote the  $\sigma$ -field of  $\mathfrak{S}$ -invariant subsets of  $A$ . Fix  $B \in \mathcal{I}$ , and define

$$\phi : \Omega_\downarrow \rightarrow [0, 1], \quad \phi(\underline{T}) \equiv \nu_A(\underline{T} \in B | T_0 = \underline{T}).$$

The  $\mathfrak{S}$ -invariance of  $B$  together with the Markov property implies

$$\nu_A[\underline{T} \in B | \mathcal{H}_i] = \nu_A[\mathfrak{S}^i \underline{T} \in B | \mathcal{H}_i] = \nu_A[\mathfrak{S}^i \underline{T} \in B | T_{n_A^i}] = \phi(T_{n_A^i}),$$

i.e.,  $\phi(T_{n_A^i})$  is an  $(\mathcal{H}_i)$ -martingale. By Lévy's upward theorem,  $\lim_{i \rightarrow \infty} \phi(T_{n_A^i}) = \mathbf{1}_B$ ,  $\nu_A$ -a.s., so that for any  $0 < a \leq b < 1$ ,

$$\frac{1}{t} \sum_{i=0}^{t-1} \mathbf{1}_{\{\phi(T_{n_A^i}) \in [a, b]\}}$$

converges  $\nu_A$ -a.s. to zero. On the other hand, by the Birkhoff ergodic theorem (see e.g. [10, Thm. 6.2.1]), it converges  $\nu_A$ -a.s. to  $\nu_A(\phi(\mathcal{T}) \in [a, b] | \mathcal{I})$ . Taking expectations on both sides we find  $\phi(\mathcal{T}) \in \{0, 1\}$   $\nu_A$ -a.s., that is,  $\phi = \mathbf{1}_{C_0}$  for some  $C_0 \in \mathcal{B}_{\Omega_\downarrow}$ . Further, since  $\phi$  is a  $\{0, 1\}$ -valued martingale, it holds  $\nu$ -a.s. that  $T_0 \in C_0$  if and only if  $T_{n_A} \in C_0$ . Since  $\nu_A(T_{n_A} = \mathfrak{S}T_0) > 0$  where  $\mathfrak{S}$  is as defined above,  $\mathbf{1}_{C_0} \leq \mathbf{1}_{\mathfrak{S}^{-1}C_0}$   $\nu_A$ -a.s. Applying the same argument with the martingale  $1 - \phi$  gives  $\mathbf{1}_{C_0} = \mathbf{1}_{\mathfrak{S}^{-1}C_0}$   $\nu_A$ -a.s., i.e., that  $C_0$  is  $\mathcal{S}$ -invariant.

*Step 3: IMGWR-triviality of  $\mathcal{S}$ -invariant sets.*

It follows from  $B \subseteq A$  that  $C_0$  is a subset of  $A_0 \equiv \{(\mathcal{T}, \xi) \in \Omega_\downarrow : |\mathcal{T}^{(o)}| = \infty\}$ . Since  $\text{IMGWR} \ll \text{IMGW}_1$  on  $A_0$  by (2.8), the result follows by showing that  $\mathcal{S}$ -invariant subsets of  $A_0$  are  $\text{IMGW}_1$ -trivial. For any  $C'_0 \subseteq A_0$  let  $\underline{C}'_0 \equiv \{(\mathcal{T}, \bar{\xi}) : (\mathcal{T}, \bar{\xi}_-) \in C'_0\}$ ; the  $\mathcal{S}$ -invariance of  $C'_0$  implies  $\underline{\mathcal{S}}$ -invariance of  $\underline{C}'_0$ . But the ergodicity of the Markov chain of types along the line  $\bar{\xi}$  readily implies that  $\underline{\mathcal{S}}$ -invariant subsets of  $\underline{A}_0$  are  $\text{IMGW}$ -trivial, e.g. by the following modification of the argument of [17, Thm. 2.15]: take  $\underline{C}'_0$  measurable with respect to the portion  $\bar{\xi}_{[-n, 0]}$  of the line between  $\bar{\xi}_{-n}$  and



$o$ , together with the descendant subtrees of  $\bar{\xi}_{-n}, \dots, \bar{\xi}_{-1}$  away from  $\bar{\xi}$ , such that the symmetric difference  $\underline{C}_0^m \triangle \underline{C}_0$  has  $\underline{\text{IMGW}}$ -measure tending to zero in  $n$ . It follows from  $\underline{\mathcal{S}}$ -invariance of  $\underline{C}_0$  together with  $\underline{\mathcal{S}}$ -stationarity of  $\underline{\text{IMGW}}$  that  $\underline{\text{IMGW}}[\underline{C}_0^m \triangle \underline{C}_0] = \underline{\text{IMGW}}[(\underline{\mathcal{S}}^m \underline{C}_0) \triangle \underline{C}_0]$  for any  $m$ , so (by the triangle inequality)

$$\lim_{n \rightarrow \infty} \sup_m |\underline{\text{IMGW}}(\underline{C}_0) - \underline{\text{IMGW}}[\underline{C}_0^m \cap (\underline{\mathcal{S}}^m \underline{C}_0)]| = 0.$$

But for any  $m > n$  we have

$$\underline{\text{IMGW}}[\underline{C}_0^m \cap (\underline{\mathcal{S}}^m \underline{C}_0)] = \underline{\text{IMGW}}[\bar{\xi}_{[-n,0]}, \bar{\xi}_{[m-n,m]}] \underline{\text{IMGW}}[\underline{C}_0^m | \bar{\xi}_{[-n,0]}] \underline{\text{IMGW}}[\underline{\mathcal{S}}^m \underline{C}_0^m | \bar{\xi}_{[m-n,m]}],$$

which tends as  $m \rightarrow \infty$  to  $\underline{\text{IMGW}}[\underline{C}_0^m]^2$ . Therefore

$$\underline{\text{IMGW}}[\underline{C}_0] = \lim_{n \rightarrow \infty} \underline{\text{IMGW}}[\underline{C}_0^n] = \lim_{n \rightarrow \infty} \underline{\text{IMGW}}[\underline{C}_0^n]^2 = \underline{\text{IMGW}}[\underline{C}_0]^2$$

which gives  $\underline{\text{IMGW}}[\underline{C}_0] \in \{0, 1\}$  as required.  $\square$

### 3. HARMONIC COORDINATES AND QUENCHED IMGWR-CLT

In this section we prove the quenched IMGWR-CLT Thm. 1.5. Let

$$\eta \equiv \mathbb{E}_{\mathbf{Q}_\infty}[W_o] = \frac{\mathbb{E}_{\text{MGW}}[W_o^2]}{\mathbb{E}_{\mathbf{g}}[e_\chi]}, \quad \sigma^2 \equiv \frac{\mathbb{E}_{\mathbf{g}}[e_\chi]^2}{\mathbb{E}_{\text{MGW}}[W_o^2]}. \quad (3.1)$$

In §3.1 we construct harmonic coordinates for  $\text{RW}_\rho$  on IMGWR-a.e.  $(\mathcal{T}, \xi)$ , and use the ergodicity result Thm. 1.4 (b) proved above to show an IMGWR-a.s. CLT for the martingale  $M_t \equiv S_{Y_t}$ , with  $M_{[nt]}/(\eta\sigma\sqrt{n})$  converging to standard Brownian motion. In §3.2 we control the error between  $h(Y_t)$  and  $M_t/\eta$  to prove Thm. 1.5. The following result, whose proof is deferred to §6.2, implies finiteness of  $\eta$  and  $\sigma$  under (H3<sup>2</sup>):

**Proposition 3.1.** *If (H1), (H2), and (H3<sup>p</sup>) hold with  $p > 1$ , then  $\mathbb{E}_{\text{MGW}}[W_o^p] < \infty$ .*

**3.1. Harmonic coordinates for  $\text{RW}_\rho$  and martingale CLT.** From now on, if  $\mu$  is a probability measure on trees (with or without marked ray), we use  $\mu$  as shorthand also for  $\mu \otimes \text{RW}_\rho$ . We write  $\mathbb{P}_{\mathcal{T}}$  for the law of the quenched random walk  $\text{RW}_\rho(\mathcal{T})$  and  $\mathbb{E}_{\mathcal{T}}$  for expectation with respect to  $\mathbb{P}_{\mathcal{T}}$ , and let  $(\mathcal{G}_t^{\mathcal{T}})_{t \geq 0}$  denote the corresponding filtration of the walk. Given  $\mathcal{T}$ , for a vertex  $v \in \mathcal{T}$  we let  $\partial v$  denote the neighbors of  $v$ , and  $\partial^+ v$  the offspring of  $v$ , i.e.,  $\partial^+ v = \partial v \cap \mathcal{T}^{(v)}$ . We write  $v \leq w$  if  $w \in \mathcal{T}^{(v)}$ , with  $v < w$  if  $w \neq v$ .

For  $v \in \mathcal{T}$  recall that  $W_v$  denotes the normalized population size of the subtree  $\mathcal{T}^{(v)}$ .<sup>1</sup> For vertices  $v \in \mathcal{T}$  we define  $S_v$  as in [31, §3]: if  $\mathcal{T}$  is a rooted tree, let

$$S_v \equiv \sum_{o < u \leq v} W_u. \quad (3.2)$$

If  $\mathcal{T}$  has marked ray  $\xi$ , recalling (1.1) we set

$$S_v \equiv S_{R_v} + S_v^\xi \quad \text{where} \quad S_{R_v} \equiv - \sum_{u \in \xi, o \geq u > R_v} W_u, \quad S_v^\xi \equiv \sum_{R_v < u \leq v} W_u. \quad (3.3)$$

<sup>1</sup>Note that if  $\mathcal{T}$  has a marked ray  $\xi$ , then for  $v \in \xi$ ,  $\mathfrak{Z}_n^v = \langle \underline{Z}_n^v, \underline{e} \rangle / \rho^n$  is not necessarily a martingale for the first  $|h(v)|$  steps. Nevertheless it is eventually a martingale so we can still define  $W_v$  to be the a.s. limit of  $\mathfrak{Z}_n^v$ .

While on MGW-a.e.  $\mathcal{T}$  the map  $v \mapsto S_v$  is harmonic except at  $o$  with respect to the transition probabilities of  $\text{RW}_\rho(\mathcal{T})$ , on IMGW-a.e.  $(\mathcal{T}, \xi)$  the map  $v \mapsto S_v$  is harmonic at every vertex with respect to the transition probabilities of  $\text{RW}_\rho(\mathcal{T}, \xi)$ . Thus, if  $(Y_t)_{t \geq 0} \sim \text{RW}_\rho(\mathcal{T}, \xi)$ ,  $M_t \equiv S_{Y_t}$  will be a martingale given a fixed realization of the tree; we regard it as providing “harmonic coordinates” for the random walk. Using the reversing measure IMGWR it is easy to prove a quenched CLT for  $M$  (extending [31, Cor. 1]):

**Proposition 3.2.** *Under (H1), (H2), and (H3<sup>2</sup>), on IMGW-a.e.  $(\mathcal{T}, \xi)$  the process  $M_{\lfloor nt \rfloor} / (\eta \sigma \sqrt{n})$  converges in distribution to a standard Brownian motion as  $n \rightarrow \infty$ .*

*Proof.* We check the conditions of the Lindeberg–Feller martingale CLT (see e.g. [10, Thm. 7.7.4]): letting

$$V_n = \frac{1}{n} \sum_{t=0}^{n-1} \mathbb{E}_{\mathcal{T}}[(M_{t+1} - M_t)^2 | \mathcal{G}_t^{\mathcal{T}}],$$

we verify that for IMGW-a.e.  $(\mathcal{T}, \xi)$ ,

- (i)  $V_n \rightarrow \eta^2 \sigma^2$  in probability and
- (ii) for all  $\epsilon > 0$ ,  $\frac{1}{n} \sum_{t=0}^{n-1} \mathbb{E}_{\mathcal{T}}[(M_{t+1} - M_t)^2 \mathbf{1}_{\{|M_{t+1} - M_t| > \epsilon \sqrt{n}\}}] \rightarrow 0$ .

Let  $Y_n$  denote the random walk on  $(\mathcal{T}, \xi)$ : we rewrite  $V_n$  in terms of the induced random walk on  $\Omega_{\downarrow}$  as

$$V_n = \frac{1}{n} \sum_{t=1}^{n-1} \varphi[(\mathcal{T}, \xi)^{Y_t}], \quad \varphi[(\mathcal{T}, \xi)] \equiv \frac{\rho}{\rho + d_o} W_o^2 + \frac{1}{\rho + d_o} \sum_{j=1}^{d_o} W_{0j}^2.$$

By Thm. 1.4 (b) and the Birkhoff ergodic theorem, we have  $V_n$  converging IMGWR-a.s. to  $\mathbb{E}_{\text{IMGWR}}[\varphi]$  provided  $\varphi \in L^1(\text{IMGWR})$ . We calculate

$$\mathbb{E}_{\text{IMGWR}}[\varphi] = \frac{1}{2\rho} \mathbb{E}_{\text{MGW}} \left[ \rho W_o^2 + \sum_{v \in \partial o} W_v^2 \right] = \mathbb{E}_{\text{MGW}}[W_o^2] = \eta^2 \sigma^2,$$

so condition (i) is proved. Condition (ii) is checked similarly using dominated convergence.  $\square$

**Remark 3.3.** To give some indication of how our results might be extended to  $\text{RWRE}_{\bar{\rho}}$ , we note that the main ingredient needed is the appropriate generalization of the normalized population size: we define it to be the random variable  $\bar{W}_o$  which is the a.s. limit of the martingale  $\mathfrak{Z}_n \equiv \mathfrak{Z}_n^{(1)}$  defined by (5.1). If  $\bar{W}_v$  denotes the normalized population size of  $\mathcal{T}^{(v)}$ , then

$$\bar{\rho} \bar{W}_v = \sum_{w \in \partial^+ v} \alpha_w \bar{W}_w,$$

so the  $\bar{W}_v$  can be used to define harmonic coordinates for the RWRE. In the single-type case,  $\bar{W}_o$  has finite second moment if and only if  $\kappa \geq 2$  [23, Thm. 2.1], so clearly Propn. 3.2 cannot apply outside this regime. We emphasize again that due to the same technical barriers which arise in [12], simple adaptations of our proof will not cover the full regime  $\kappa \geq 2$ .

**3.2. Quenched IMGWR-CLT.** We now prove the quenched CLT for IMGWR trees by controlling the corrector

$$\varepsilon_t \equiv \frac{M_t}{\eta} - h(Y_t)$$

on the interval  $0 \leq t \leq n$ . For  $1/2 < \delta < 1$  and  $n \geq 0$  fixed, let  $\tau^n(j)$ , for  $j \lfloor n^\delta \rfloor \leq n$  denote integer times chosen uniformly at random (independently of one another and of the random walk  $Y$ ) from the interval  $[j \lfloor n^\delta \rfloor, (j+1) \lfloor n^\delta \rfloor]$ .

**Proposition 3.4.** *Assume (H1), (H2), and (H3<sup>p</sup>) with  $p > 2$ . There exists  $\delta_0 \equiv \delta_0(p) \in (1/2, 1)$  such that for  $\delta_0 \leq \delta < 1$  and  $\epsilon > 0$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P}_{(\mathcal{T}, \xi)} \left( \max_{j \lfloor n^\delta \rfloor \leq n} |\varepsilon_{\tau^n(j)}| \geq \epsilon \sqrt{n} \right) = 0, \quad \text{IMGWR-a.s.} \quad (3.4)$$

Further, for any  $\epsilon'$  with  $2\epsilon' + \delta < 1$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P}_{(\mathcal{T}, \xi)} \left( \max_{r, s \leq n, |r-s| \leq n^\delta} |h(Y_r) - h(Y_s)| \geq n^{1/2-\epsilon'} \right) = 0, \quad \text{IMGWR-a.s.} \quad (3.5)$$

Given this proposition, we can prove the quenched CLT for  $\text{RW}_\rho$  on IMGWR trees:

*Proof of Thm. 1.5.* If  $t \leq n$  then  $|t - \tau^n(j)| \leq \lfloor n^\delta \rfloor$  for some  $j$ , so

$$\max_{t \leq n} |\varepsilon_t| \leq \max_{r, s \leq n, |r-s| \leq \lfloor n^\delta \rfloor} \left| \frac{M_r}{\eta} - \frac{M_s}{\eta} \right| + \max_{j \lfloor n^\delta \rfloor \leq n} |\varepsilon_{\tau^n(j)}| + \max_{r, s \leq n, |r-s| \leq \lfloor n^\delta \rfloor} |h(Y_r) - h(Y_s)|.$$

$M$  satisfies a CLT by Propn. 3.2, and it follows from (3.4) and (3.5) that

$$\lim_{n \rightarrow \infty} \mathbb{P}_{(\mathcal{T}, \xi)} \left( \max_{t \leq n} |\varepsilon_t| \geq \epsilon \sqrt{n} \right) = 0, \quad \text{IMGWR-a.s.,}$$

which gives the result.  $\square$

The remainder of this section is devoted to the proof of Propn. 3.4.

**3.2.1. Tightness.** We begin by proving (3.5), using some *a priori* (annealed) estimates for  $\text{RW}_\rho$  coming from the Carne–Varopoulos bound.

**Lemma 3.5.** *There exists a constant  $C < \infty$  such that*

$$\text{MGW} \left( \max_{t \leq n} |X_t| \geq m \right) \leq C n e^{-(m+1)^2/(2n)} \quad \forall m, n \geq 1.$$

*Proof.* We modify the proof of [31, Lem. 5]. Take the finite tree with vertices  $\{w \in \mathcal{T} : |w| \leq m\}$ , and make this into a wired tree  $\mathcal{T}^*$  by adding a new vertex  $o^*$  which is joined by an edge to each vertex in  $D_m$ . Define the modified random walk  $X^*$  on  $\mathcal{T}^*$  which follows the law of  $\text{RW}_\rho$  except at  $o^*$  where it moves to a vertex chosen uniformly at random from  $D_m$ . Then

$$\mathbb{P}_{\mathcal{T}}(\max_{t \leq n} |X_t| \geq m) \leq 2 \sum_{t=1}^{n+1} \mathbb{P}_{\mathcal{T}^*}(X_t = o^*).$$

By the Carne–Varopoulos inequality (see [28, Thm. 13.4]),

$$\mathbb{P}_{\mathcal{T}^*}(X_t^* = o^*) \leq 2\sqrt{\frac{|D_m|}{\rho^{m-1}}}e^{-(m+1)^2/(2t)}.$$

Taking expectations gives

$$\text{MGW}(|X_t^*| = o^*) \leq Ce^{-(m+1)^2/(2t)},$$

and summing over  $1 \leq t \leq n+1$  gives the result.  $\square$

**Corollary 3.6.** *There exists a constant  $C < \infty$  such that for any  $m, n \geq 1$ ,*

$$C^{-1} \text{IMGW}_0\left(\max_{t \leq n} |h(Y_t)| \geq m\right) \leq \text{IMGWR}\left(\max_{t \leq n} |h(Y_t)| \geq m\right) \leq Cn^2 e^{-m^2/(2n)}.$$

*Proof.* We argue as in the proof of [31, Cor. 2]. By decomposing into at most  $n$  excursions away from height zero and using the stationarity of IMGWR, we find

$$\begin{aligned} & \text{IMGWR}\left(\max_{t \leq n} h(Y_t) \geq m\right) \\ & \leq n \text{IMGWR}\left(\exists t \leq n : h(Y_t) \geq m, h(Y_s) > 0 \forall 0 \leq s \leq t\right) \\ & \leq Cn \text{MGW}\left(\max_{t \leq n} |X_t| \geq m-1\right) \leq Cn^2 e^{-m^2/(2n)}, \end{aligned}$$

by Lem. 3.5. The same bound holds for  $\text{IMGWR}(\min_{t \leq n} h(Y_t) \leq -m)$  by the reversibility of IMGWR. The result follows by noting that  $d\text{IMGW}_0/d\text{IMGWR}$  is uniformly bounded by a deterministic constant.  $\square$

*Proof of Propn. 3.4, (3.5).* By stationarity of IMGWR and Cor. 3.6, for any fixed  $s$

$$\text{IMGWR}\left(\max_{0 \leq u \leq n^\delta} |h(Y_{s+u}) - h(Y_s)| \geq n^{1/2-\epsilon'}\right) \leq Cn^{2\delta} e^{-n^{1-2\epsilon'-\delta}/2},$$

and summing over  $s \leq n$  gives

$$\text{IMGWR}\left(\max_{r, s \leq t, |r-s| \leq n^\delta} |h(Y_r) - h(Y_s)| \geq n^{1/2-\epsilon'}\right) \leq Cn^{2\delta+1} e^{-n^{1-2\epsilon'-\delta}/2},$$

which is summable in  $n$  provided  $2\epsilon' + \delta < 1$ . The result then follows from Markov's inequality and Borel–Cantelli.  $\square$

**3.2.2. Control of corrector.** In the remainder of this section we prove (3.4). We will make use of the following classical result:

**Lemma 3.7** ([33, p. 60]). *If  $z_1, \dots, z_n$  are independent random variables with  $\mathbb{E}z_i = 0$  and  $\mathbb{E}|z_i|^p < \infty$ , then*

$$\mathbb{E}\left[\left|\sum_{i=1}^n z_i\right|^p\right] \leq \begin{cases} 2 \sum_{i=1}^n \mathbb{E}[|z_i|^p] & \text{if } 1 \leq p \leq 2, \\ C(p)n^{p/2-1} \sum_{i=1}^n \mathbb{E}[|z_i|^p] & \text{if } p \geq 2. \end{cases}$$

Recalling (1.1) and (3.3), we decompose

$$\frac{1}{\sqrt{n}} \max_{j \lfloor n^\delta \rfloor \leq n} |\varepsilon_{\tau^n(j)}| \leq E_1 + E_2 \quad (3.6)$$

where, with  $R_t \equiv R_{Y_t}$  denoting the nearest ancestor of  $Y_t$  on  $\xi$ ,

$$E_1 \equiv \frac{1}{\sqrt{n}} \max_{t \leq 2n} \left| \frac{S_{R_t}}{\eta} - h(R_t) \right|, \quad E_2 \equiv \frac{1}{\sqrt{n}} \max_{j \lfloor n^\delta \rfloor \leq n} \left| \frac{S_{Y_{\tau^n(j)}}^\xi}{\eta} - d(Y_{\tau^n(j)}, \xi) \right|.$$

The following lemma says that the harmonic coordinates  $(S_v)_{v \in \mathcal{T}}$  of (3.2), rescaled by  $\eta$  of (3.1), are a good approximation to the actual coordinates  $|v|$  on the **MGW** rooted trees. Let

$$A_n^\epsilon \equiv A_n^\epsilon(\mathcal{T}) \equiv \left\{ v \in D_n : \left| \frac{S_v}{n} - \eta \right| > \epsilon \right\}, \quad \epsilon > 0, n \geq 1. \quad (3.7)$$

Let  $\tau^\circ \equiv \min\{t > 0 : X_t = o\}$  denote the first return time to the starting point  $X_0 = o$  by the walk  $X$ .

**Lemma 3.8.** *Assume (H1), (H2), and (H3<sup>p</sup>) with  $p \geq 2$ . For any  $\epsilon > 0$ , the expected number of visits to  $A_k^\epsilon$  during a single excursion away from the root is*

$$\mathbb{E}_{\text{MGW}} \left[ \sum_{t=0}^{\tau^\circ} \mathbf{1}_{\{X_t \in A_k^\epsilon\}} \right] \leq \frac{C}{\rho^k} \mathbb{E}_{\text{MGW}} \left[ \sum_{v \in A_k^\epsilon} (1 + d_v) \right] \leq \frac{C(p, \epsilon)}{k^{p/2}}.$$

*Proof.* If  $v \in \mathcal{T}$  with  $|v| = k \geq 1$ , a simple conductance calculation (see [28, Ch. 2]) gives

$$\mathbb{E}_{\mathcal{T}} \left[ \sum_{t=0}^{\tau^\circ} \mathbf{1}_{\{X_t=v\}} \right] = \frac{\mathbb{P}_{\mathcal{T}}(o \rightarrow v)}{\mathbb{P}_{\mathcal{T}}(v \rightarrow o)} = \frac{\rho + d_v}{d_o \rho^k}, \quad (3.8)$$

so the first inequality follows. For the second we follow the proof of [31, Lem. 3] (in particular the estimate [31, (20)]) and of [12, Lem. 4.2]. Recall from §2.3 the definition (2.7) of the probability measure  $\mathbf{Q}_k^a$  on rooted trees  $\mathcal{T}$  given by a size-biasing of  $\text{MGW}^a$ , and further the probability  $\mathbf{Q}_{k*}^a$  on rooted trees  $\mathcal{T}$  with a marked path  $(o = v_0, \dots, v_k)$  from the root to level  $k$ :

$$\begin{aligned} \mathbb{E}_{\text{MGW}^a} \left[ \sum_{v \in A_k^\epsilon} (1 + d_v) \right] &\leq C \mathbb{E}_{\text{MGW}^a} \left[ \sum_{v \in A_k^\epsilon} e_{\chi_v} (1 + d_v) \right] \\ &\leq C \rho^k \mathbb{E}_{\mathbf{Q}^a} \left[ \frac{\sum_{v \in A_k^\epsilon} e_{\chi_v} (1 + d_v)}{\langle \underline{Z}_k, \underline{e} \rangle} \right] = C \rho^k \mathbb{E}_{\mathbf{Q}_{k*}^a} [(1 + d_{v_k}) \mathbf{1}_{\{v_k \in A_k^\epsilon\}}], \end{aligned}$$

so it suffices to show

$$\mathbb{E}_{\mathbf{Q}_{k*}^a} [(1 + d_{v_k}) \mathbf{1}_{\{v_k \in A_k^\epsilon\}}] \leq C(p, \epsilon) k^{-p/2}.$$

To this end, writing  $W_i \equiv W_{v_i}$ , for  $i < k$  we decompose  $W_i \equiv W_{i+1}/\rho + W_i^\lambda$  where  $W_i^\lambda$  is the normalized population size of  $\mathcal{T}^{(v_i)} \setminus \mathcal{T}^{(v_{i+1})}$ . Then

$$W_i = \sum_{j=i}^{k-1} \frac{W_j^\lambda}{\rho^{j-i}} + \frac{W_k}{\rho^{k-i}},$$

so

$$\frac{S_{v_k}}{k} - \eta = \frac{1}{k} C_k W_k + \frac{1}{k} \sum_{i=1}^{k-1} C_i W_i^\lambda - \eta, \quad C_i \equiv \sum_{j=0}^{i-1} \rho^{-j} \leq C_\infty \equiv \frac{\rho}{\rho-1}.$$

Conditional on the types  $(\chi_i \equiv \chi_{v_i})_{i=1}^k$ , the random variables  $W_1^\lambda, \dots, W_{k-1}^\lambda$  are independent of one another and of the pair  $(W_k, d_{v_k})$ , and all these random variables have finite moments of order  $p$  by Propn. 3.1. Therefore

$$\mathbb{E}_{\mathbf{Q}_{k*}^a}[(1+d_{v_k})\mathbf{1}_{\{v_k \in A_k^\epsilon\}}] \leq C\mathbf{Q}_{k*}^a\left(\left|\frac{1}{k} \sum_{i=1}^{k-1} C_i W_i^\lambda - \eta\right| \geq \epsilon/2\right) + \mathbb{E}_{\mathbf{Q}_{k*}^a}[(1+d_{v_k})\mathbf{1}_{\{C_k W_k/k \geq \epsilon/2\}}]$$

By (H3<sup>p</sup>), Markov's inequality, and Hölder's inequality, the second term is

$$\leq \left(\frac{2C_k}{k\epsilon}\right)^{p-1} \mathbb{E}_{\mathbf{Q}_{k*}^a}[(1+d_{v_k})W_k^{p-1}] \leq \frac{C(p, \epsilon)}{k^{p-1}} \leq \frac{C(p, \epsilon)}{k^{p/2}}$$

(since  $p \geq 2$ ). As for the first term, by Lem. 3.7 and Markov's inequality,

$$\begin{aligned} & \mathbf{Q}_{k*}^a\left(\left|\frac{1}{k} \sum_{i=1}^{k-1} C_i W_i^\lambda - \mathbb{E}_{\mathbf{Q}_{k*}^a}\left[\frac{1}{k} \sum_{i=1}^{k-1} C_i W_i^\lambda \mid (\chi_i)_{i=1}^k\right]\right| > \epsilon/4\right) \\ & \leq C(p) \frac{k^{p/2-1}}{(k\epsilon)^p} \left\{ \sum_{i=1}^{k-1} \mathbb{E}_{\mathbf{Q}_{k*}^a}[|C_i(W_i^\lambda - \mathbb{E}[W_i^\lambda \mid \chi_i])|^p] \right\} \leq C(p, \epsilon) k^{-p/2}. \end{aligned}$$

On the other hand,

$$\mathbb{E}_{\mathbf{Q}_{k*}^a}\left[\frac{1}{k} \sum_{i=1}^{k-1} C_i W_i^\lambda\right] \rightarrow C_\infty \mathbb{E}_{\mathbf{Q}_\infty}[W_o^\lambda] = \mathbb{E}_{\mathbf{Q}_\infty}[W_o] = \eta,$$

and so

$$\mathbf{Q}_{k*}^a\left(\left|\mathbb{E}_{\mathbf{Q}_{k*}^a}\left[\frac{1}{k} \sum_{i=1}^{k-1} C_i W_i^\lambda \mid (\chi_i)_{i=1}^k\right] - \eta\right| > \epsilon/4\right)$$

decays exponentially in  $k$  by [9, Thm. 3.1.2]. Combining these estimates completes the proof.  $\square$

Recalling the definition (3.3) of the harmonic coordinates on the IMGWR trees, the next step is to use Lem. 3.8 to show that on these trees  $S_v^\xi/\eta$  is a good approximation to  $d(v, \xi)$ . In analogy with (3.7) set

$$B_k^\epsilon = \left\{ w \in \mathcal{T} : d(w, \xi) = k, \left| \frac{S_w^\xi}{k} - \eta \right| > \epsilon \right\}, \quad B^\epsilon = \bigcup_{k \geq 1} B_k^\epsilon(\mathcal{T}, \xi). \quad (3.9)$$

**Lemma 3.9.** *Assume (H1), (H2), and (H3<sup>p</sup>) with  $p > 2$ . There exists  $\delta_0 \equiv \delta_0(p) \in (1/2, 1)$  such that for  $\delta_0 \leq \delta < 1$  and  $\epsilon > 0$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P}_{(\mathcal{T}, \xi)} \left( \exists j \in \mathbb{Z}_{\geq 0}, j \lfloor n^\delta \rfloor \leq n : Y_{\tau^n(j)} \in B^\epsilon \right) = 0, \quad \text{IMGWR-a.s.}$$

*Proof.* We modify the proof of [12, (22)]. If we define

$$\tau_{n,\epsilon}^{\text{hit}} \equiv \inf \{ t \geq 0 : |h(Y_t)| = \lfloor n^{1/2+\epsilon} \rfloor \},$$

then Cor. 3.6 together with Markov's inequality gives

$$\text{IMGWR}[\mathbb{P}_{(\mathcal{T}, \xi)}(\tau_{n,\epsilon}^{\text{hit}} \leq n) \geq c] \leq c^{-1} \text{IMGWR}(\tau_{n,\epsilon}^{\text{hit}} \leq n) \leq c^{-1} C n^2 e^{-n^{2\epsilon/2}},$$

so by Borel–Cantelli we have  $\mathbb{P}_{(\mathcal{T}, \xi)}(\tau_{n,\epsilon}^{\text{hit}} \leq n) \rightarrow 0$ , IMGWR-a.s.

On the event  $\{\tau_{n,\epsilon}^{\text{hit}} > n\}$ , we decompose the walk into excursions from  $\xi$  started at  $v_i$ ,  $0 \leq i < \lfloor n^{1/2+\epsilon} \rfloor$  (with each step of the walk along the ray contributing an empty excursion) and apply Wald's identity (see e.g. [3, Exercise 22.8]) to find

$$\begin{aligned} & \mathbb{P}_{(\mathcal{T}, \xi)} \left( \left\{ \exists j \lfloor n^\delta \rfloor \leq n : Y_{\tau^n(j)} \in B^\epsilon \right\} \cap \{ \tau_{n,\epsilon}^{\text{hit}} > n \} \right) \\ & \leq \frac{1}{\lfloor n^\delta \rfloor} \sum_{i=0}^{\lfloor n^{1/2+\epsilon} \rfloor - 1} \mathbb{E}_{(\mathcal{T}, \xi)}[L_i(\tau_{n,\epsilon}^{\text{hit}})] \mathbb{E}_{(\mathcal{T}, \xi)}^i[L(B^\epsilon; \tau^{\text{exc}})]. \end{aligned} \quad (3.10)$$

In the above,  $L_A(n) \equiv L(A; n)$  denotes the number of visits to set  $A$  by time  $n$  and  $L_i(n) \equiv L(v_i; n)$ .  $\mathbb{E}_{(\mathcal{T}, \xi)}^i$  denotes expectation with respect to the law of a  $\rho$ -biased random walk  $Y$  started from  $Y_0 = v_i$ , and  $\tau^{\text{exc}} \equiv \inf\{t > 0 : Y_t = v_i \text{ or } Y_t \notin \mathcal{T}^{(v_i)}\}$  denotes the excursion end time.

By a conductance calculation,

$$\mathbb{E}_{(\mathcal{T}, \xi)}[L_i(\tau_{n,\epsilon}^{\text{hit}})] = \frac{1}{\mathbb{P}(v_i \rightarrow v_{\lfloor n^{1+\epsilon} \rfloor})} \leq \frac{1 + d_i/\rho}{1 - \rho} \leq C d_i. \quad (3.11)$$

During a single excursion away from  $\xi$  the walk can visit only one of the  $\mathcal{T}^{(w)}$  for  $w \in \partial^+ v \setminus v_{i-1}$ , so to bound the second factor of each summand in (3.10) it suffices to consider an MGW rooted tree  $\mathcal{T}'$  (without ray): letting

$$\tilde{A}_k^\epsilon \equiv \tilde{A}_k^\epsilon(\mathcal{T}') \equiv \left\{ v \in D_k : \left| \frac{W_o + S_v}{k+1} - \eta \right| > \epsilon \right\}, \quad \tilde{A}^\epsilon \equiv \bigcup_{k \geq 0} \tilde{A}_k^\epsilon,$$

it follows from a (very slight) modification of Lem. 3.8 that

$$\mathbb{E}_{\text{IMGW}_0}[L(B^\epsilon; \tau^{\text{exc}}) | i \cup \partial^+ i] \leq C \mathbb{E}_{\text{MGW}} \left[ \sum_{t=0}^{\tau^\circ} \mathbf{1}_{\{X_t \in \tilde{A}^\epsilon\}} \right] \leq C(p, \epsilon) \sum_{k \geq 1} k^{-p/2} \leq C(p, \epsilon)$$

(using  $p > 2$ ). It follows that the quantity in (3.10) converges to zero IMGWR-a.s., which concludes the proof.  $\square$



*Proof of Propn. 3.4, (3.4).* Recall the decomposition (3.6). For any  $k_0$ ,

$$E_1 \leq \frac{1}{\sqrt{n}} \max_{i \leq k_0} \left| \frac{S_{v_i}}{\eta} - h(v_i) \right| + \frac{1}{\sqrt{n}} \max_{t \leq 2n, h(R_t) > k_0} \left| \frac{S_{R_t}}{\eta} - h(R_t) \right|.$$

The first term clearly tends to zero as  $n \rightarrow \infty$  with  $k_0$  fixed. The second term is bounded above by

$$\left( \frac{1}{\sqrt{n}} \max_{t \leq 2n} |M_t| \right) \sup_{i > k_0} \left| \frac{1}{\eta} - \frac{h(v_i)}{S_{v_i}} \right|. \quad (3.12)$$

Now recall from the proof of Propn. 2.1 that if  $(\mathcal{T}, \xi) \sim \text{IMGW}_0$  then  $(\mathcal{T}^{(k)}, o) \sim \mathbf{Q}_{k\star}$ . Thus a consequence of the proof of Lem. 3.8 is that for sufficiently small  $\epsilon$ ,

$$\text{IMGW}_0(|S_{v_k}/k + \eta| \geq \epsilon) \leq C(p, \epsilon)k^{-p/2}.$$

Therefore the supremum in (3.12) can be made arbitrarily small by taking  $k_0$  large. We also have

$$E_2 \leq \left( \frac{1}{\sqrt{n}} \max_{t \leq 2n} |M_t| \right) \max_{j \lfloor n^\delta \rfloor \leq n} \left| \frac{1}{\eta} - \frac{d(Y_{\tau^n(j)}, \xi)}{S_{Y_{\tau^n(j)}}^\xi} \right|,$$

and in view of Lem. 3.9 the second factor tends to zero in probability. By the invariance principle for  $M$  proved in Propn. 3.2,  $\max_{t \leq 2n} |M_t|/\sqrt{n}$  stays bounded in probability as  $n \rightarrow \infty$ , so the result follows.  $\square$

#### 4. FROM IMGWR-CLT TO MGW-CLT BY SHIFTED COUPLING

In this section we prove our main result Thm. 1.1. In §4.1 we review (a slight modification of) the “shifted coupling” procedure of [31, §6], which we use in §4.2 to transfer the IMGWR-CLT to an annealed MGW-CLT. In §4.3 we prove a variance estimate which allows to go from the annealed to the quenched MGW-CLT.

**4.1. The shifted coupling construction.** We begin by reviewing the shifted coupling construction of [31, §6], with the (natural) modification needed to handle the multi-type case. The basic observation underlying the construction is that the law of the random walk  $X \sim \text{RW}_\rho(\mathcal{T})$  up to time  $t$  depends only on

$$\mathcal{E}_t \equiv o \cup (\partial X_s)_{0 \leq s < t}$$

(“the subtree explored by time  $t$ ”), so that one can construct the tree at the same time as the random walk.

For any tree  $\mathcal{T}$  (with or without marked ray) and  $U$  any subset of the vertices of  $\mathcal{T}$ , we also use  $U$  to indicated the subgraph of  $\mathcal{T}$  induced by  $U$ . Let  $\mathcal{L}\mathcal{T}$  denote the set of leaves and  $\mathcal{T}^\circ \equiv \mathcal{T} \setminus \mathcal{L}\mathcal{T}$ .

Let  $a_0 \in \mathcal{Q}$  be fixed, and suppose  $(\mathcal{T}, (X_t)_{t \geq 0}) \sim \text{MGW} \otimes \text{RW}_\rho$ . For each fixed  $n \geq 1$  we give a decomposition of  $X$  into “fresh excursions” marked by time intervals  $[\tau_i, \eta_i]$ ,  $i \geq 1$ , as follows. Set  $\eta_0 \equiv 0$  and define

$$\ell(n) \equiv 4 \lfloor (\log(1+n))^{3/2} \rfloor. \quad (4.1)$$

For  $i \geq 1$ , let

$$\begin{aligned}\tau_i &\equiv \min\{t > \eta_{i-1} : X_t \in \mathcal{LE}_t, |X_t| > \ell(n)/2, \chi_{X_t} = a_0\}, && \text{excursion start,} \\ \eta_i &\equiv \min\{t > \tau_i : X_t \in \mathcal{E}_{\tau_i}^\circ\}, && \text{excursion end,} \\ \mathcal{V}_i &\equiv X_{\tau_i} \cup \mathcal{E}_{\eta_i} \setminus \mathcal{E}_{\tau_i}, && \text{excursion exploration.}\end{aligned}$$

We take the convention  $\min \emptyset \equiv \infty$ , and let  $\tau^\mathbb{X} \equiv \max\{i : \eta_i < \infty\}$  be the total number of excursions (so  $\{\tau^\mathbb{X} < \infty\} = \mathbb{X}$ ).

Next we construct a *coupled* realization  $((\mathcal{T}^\diamond, \xi), (Y_t)_{t \geq 0}) \sim \text{IMGW}_0 \otimes \text{RW}_\rho$  as follows: first construct the backbone  $\mathcal{E}_0^\diamond$  of the tree ( $\xi$  and  $\partial^+ v_i$  for  $i \geq 1$ , together with types) in the manner described in §2.1. Set  $\eta_0^\diamond \equiv 0$ , and start a  $\rho$ -biased random walk  $Y$  on  $\mathcal{E}_0^\diamond$  with  $Y_0 = o$ . As in the **MGW** setting we will construct a growing sequence  $(\mathcal{E}_t^\diamond)_{t \geq 0}$  such that  $\mathcal{E}_t^\diamond = \mathcal{E}_0^\diamond \cup (\partial Y_s)_{0 \leq s < t}$ , and we will define (for  $i \geq 1$ )

$$\begin{aligned}\tau_i^\diamond &\equiv \min\{t > \eta_{i-1}^\diamond : Y_t \in \mathcal{LE}_t^\diamond, d(Y_t, \xi) > \ell(n)/2, \chi_{Y_t} = a_0\}, && \text{excursion start,} \\ \eta_i^\diamond &\equiv \min\{t > \tau_i^\diamond : Y_t \in (\mathcal{E}_{\tau_i^\diamond}^\diamond)^\circ\}, && \text{excursion end,} \\ \mathcal{V}_i^\diamond &\equiv Y_{\tau_i^\diamond} \cup \mathcal{E}_{\eta_i^\diamond}^\diamond \setminus \mathcal{E}_{\tau_i^\diamond}^\diamond, && \text{excursion exploration.}\end{aligned}$$

The difference is that we grow the sequence  $\mathcal{E}_t^\diamond$  in a manner *dependent* on  $(\mathcal{T}, X)$ , such that excursions of  $Y$  into unexplored territory (and started from  $a_0$ ) match the excursions of  $X$  defined above: formally, we couple  $(Y_s)_{\tau_i^\diamond \leq s < \eta_i^\diamond}$  with  $(X_s)_{\tau_i \leq s < \eta_i}$  such that there is a (type-preserving) isomorphism  $f_i : \mathcal{V}_i \rightarrow \mathcal{V}_i^\diamond$  with  $f_i(X_{\tau_i+s}) = Y_{\tau_i^\diamond+s}$ , and then we set  $Y_{\eta_i^\diamond}$  to be the ancestor of  $Y_{\tau_i^\diamond}$  (not necessarily of the same type as  $X_{\eta_i}$ ). Then, on the inter-excursion intervals  $\eta_{i-1}^\diamond \leq t < \tau_i^\diamond$  (for  $i \geq 1$ ),

- If  $Y_t \in (\mathcal{E}_t^\diamond)^\circ$  then generate  $Y_{t+1}$  according to the transition kernel of  $\text{RW}_\rho$  on  $\mathcal{E}_{t+1}^\diamond = \mathcal{E}_t^\diamond$ ;
- If  $Y_t \in \mathcal{LE}_t^\diamond$  with  $\chi_{Y_t} \neq a_0$ , let  $\mathcal{E}_{t+1}^\diamond$  be the enlargement of  $\mathcal{E}_t^\diamond$  obtained by attaching random offspring to  $Y_t$  according to law  $\mathbf{q}^{\chi_{Y_t}}$ , and generate  $Y_{t+1}$  according to the transition kernel of  $\text{RW}_\rho$  on  $\mathcal{E}_{t+1}^\diamond$ .

Finally, with  $\mathcal{E}_\infty^\diamond \equiv \lim_{t \rightarrow \infty} \mathcal{E}_t^\diamond$ , we define  $\mathcal{T}^\diamond$  by attaching to each vertex  $v \in \mathcal{LE}_\infty^\diamond$  an independent  $\text{MGW}^{\chi_v}$  tree. We thus obtain the following extension of [31, Lem. 8]:

**Lemma 4.1.** *If  $(\mathcal{T}, (X_t)_{t \geq 0}) \sim \text{MGW} \otimes \text{RW}_\rho$  then the marginal law of  $((\mathcal{T}^\diamond, \xi), (Y_t)_{t \geq 0})$  arising from the above construction is  $\text{IMGW}_0 \otimes \text{RW}_\rho$ .*

**Remark 4.2.** Although we suppress the parameter  $n$  from the notation, we emphasize that each  $n \geq 1$  gives rise to a different excursion decomposition, hence a different coupling between  $(\mathcal{T}, X)$  and  $((\mathcal{T}^\diamond, \xi), Y)$ .

**4.2. Annealed MGW-CLT.** We now transfer the quenched **IMGWR-CLT** to the following annealed **MGW-CLT**:

**Proposition 4.3.** *Assume (H1), (H2) and (H3<sup>p</sup>) with  $p > 4$ . If  $(\mathcal{T}, X)$  has law  $\text{MGW} \otimes \text{RW}_\rho$  conditioned on  $\mathbb{X}^c$ , then the processes  $(|X_{\lfloor nt \rfloor}|/(\sigma\sqrt{n}))_{t \geq 0}$  converge in law to the absolute value of a standard Brownian motion.*

Recall that  $R_t \equiv R_{Y_t}$  denotes the nearest ancestor of  $Y_t$  on  $\xi$ . By Thm. 1.5, for  $\text{IMGW}_0$ -a.e.  $(\mathcal{T}, \xi)$ , the process

$$\mathbb{H}_{\lfloor nt \rfloor} / (\sigma \sqrt{n}), \quad \mathbb{H}_t \equiv h(Y_t) - \min_{0 \leq s \leq t} h(Y_s) = h(Y_t) - \min_{0 \leq s \leq t} h(R_s) \geq 0 \quad (4.2)$$

converges to a Brownian motion minus its running minimum, which is the same in law as the absolute value of a Brownian motion (see e.g. [18, Thm. 3.6.17]). Thus to deduce Propn. 4.3 we need to estimate the relation between the processes  $|X_n|$  and  $\mathbb{H}_n$ . To this end, let  $\mathbf{t}, \mathbf{t}^\diamond$  be the monotone increasing bijections

$$\mathbf{t} : \mathbb{Z}_{\geq 0} \rightarrow \bigcup_{i \geq 1} [\eta_{i-1}, \tau_i), \quad \mathbf{t}^\diamond : \mathbb{Z}_{\geq 0} \rightarrow \bigcup_{i \geq 1} [\eta_{i-1}^\diamond, \tau_i^\diamond),$$

parametrizing the inter-excursion times of  $X_n$  and  $Y_n$  respectively. We make the following notations (the left column refers to the  $\text{MGW}$  tree, while the right column refers to the  $\text{IMGW}_0$  tree):

$$\begin{aligned} X_s^{\text{int}} &\equiv X_{\mathbf{t}(s)}, & Y_s^{\text{int}} &\equiv Y_{\mathbf{t}^\diamond(s)} \\ \mathcal{H}_s &\equiv \sigma(X_t : t \leq \mathbf{t}(s)) & \mathcal{H}_s^\diamond &\equiv \sigma(Y_t : t \leq \mathbf{t}^\diamond(s)), \\ J_i &\equiv \mathbf{t}^{-1}[\eta_{i-1}, \tau_i), & J_i^\diamond &\equiv (\mathbf{t}^\diamond)^{-1}[\eta_{i-1}^\diamond, \tau_i^\diamond); \\ I_n &\equiv \max \{i : \eta_{i-1} \leq n\}, & I_n^\diamond &\equiv \max \{i : \eta_{i-1}^\diamond \leq n\}; \\ \Delta_n &\equiv \sum_{i=1}^{I_n} |J_i|, & \Delta_n^\diamond &\equiv \sum_{i=1}^{I_n^\diamond} |J_i^\diamond|; \\ \Delta_n(\alpha) &\equiv \sum_{i=1}^{I_n} |\{s \in J_i : |X_s^{\text{int}}| \leq n^\alpha\}|, & \Delta_n^\diamond(\alpha) &\equiv \sum_{i=1}^{I_n^\diamond} |\{s \in J_i^\diamond : d(Y_s^{\text{int}}, \xi) \leq n^\alpha\}|. \end{aligned}$$

In words, given the walk  $X$  on the  $\text{MGW}$  tree,  $X_s^{\text{int}}$  is the “inter-excursion process” adapted to the filtration  $\mathcal{H}_s$ ,  $J_i$  is the  $i$ -th inter-excursion interval,  $I_n$  is the number of such intervals intersecting  $[0, n]$ ,  $\Delta_n$  is the total length of these intervals, and  $\Delta_n(\alpha)$  is the length of these intervals except for times spent at distance more than  $n^\alpha$  from the root. The right column defines the analogous objects for the walk on the  $\text{IMGW}_0$  tree.

**Lemma 4.4.** *Assume (H1), (H2), and (H3<sup>p</sup>) with  $p > 4$ . There exists  $\alpha_0(p) < 1/2$  such that for  $\alpha > \alpha_0(p)$ ,*

$$\left. \begin{array}{l} \text{MGW}(\Delta_n(\alpha) \neq \Delta_n \mid \mathbb{X}^c) \\ \text{IMGW}_0(\Delta_n^\diamond(\alpha) \neq \Delta_n^\diamond) \end{array} \right\} \leq n^{-c}, \quad c \equiv c(p, \alpha) > 0.$$

We will obtain the corollary below as a relatively straightforward consequence of Lem. 4.4. Let

$$\mathbb{D}_n \equiv \max_{0 \leq r \leq s \leq n} \{h(R_s) - h(R_r)\}$$

denote the maximum displacement by time  $n$  against the backward drift on  $\xi$ .

**Corollary 4.5.** *Assume (H1), (H2), and (H3<sup>p</sup>) with  $p > 4$ . Then*

(a) *There exists  $\alpha_0(p) < 1/2$  such that for  $\alpha \geq \alpha_0(p)$ ,*

$$\left. \begin{array}{l} \text{MGW}(\Delta_n \geq n^{1/2+\alpha+\epsilon} \mid \mathbb{X}^c) \\ \text{IMGW}_0(\Delta_n^\diamond \geq n^{1/2+\alpha+\epsilon}) \end{array} \right\} \leq n^{-c}, \quad c \equiv c(p, \alpha, \epsilon) > 0$$

(b) *On  $\text{IMGW}_0$ -a.e.  $(\mathcal{T}, \xi)$ ,  $\mathbb{D}_n / \sqrt{n}$  converges  $\mathbb{P}_{(\mathcal{T}, \xi)}$ -a.s. to zero.*

Assuming these results we can prove the annealed MGW-CLT:

*Proof of Propn. 4.3.* Let  $\mathbf{b} : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$  be any nondecreasing map which maps  $[\tau_i^\diamond, \eta_i^\diamond)$  bijectively onto  $[\tau_i, \eta_i)$ , and  $J_i^\diamond$  into  $J_i$ , for each  $i$ . Then for  $t \in [\tau_i^\diamond, \eta_i^\diamond)$  we have  $|X_{\mathbf{b}(t)}| - |X_{\tau_i}| = d(Y_t, \xi) - d(Y_{\tau_i^\diamond}, \xi)$ , so, recalling (1.1) and (4.2), we have

$$||X_{\mathbf{b}(t)}| - \mathbb{H}_t| = \left| |X_{\mathbf{b}(t)}| - d(Y_t, \xi) - h(R_t) + \min_{s \leq t} h(R_s) \right| \leq |X_{\tau_i}| + d(Y_{\tau_i^\diamond}, \xi) + \mathbb{D}_t.$$

If instead  $t \in J_i^\diamond$  then

$$||X_{\mathbf{b}(t)}| - \mathbb{H}_t| \leq |X_{\mathbf{b}(t)}| + |\mathbb{H}_t| \leq |X_{\mathbf{b}(t)}| + \mathbb{D}_t + d(Y_t, \xi).$$

It follows that on the event  $\{\Delta_n(\alpha) = \Delta_n\} \cap \{\Delta_n^\diamond(\alpha) = \Delta_n^\diamond\}$ ,

$$\frac{1}{\sqrt{n}} \max_{0 \leq t \leq n} ||X_{\mathbf{b}(t)}| - \mathbb{H}_t| \leq \frac{2n^\alpha + \mathbb{D}_n}{\sqrt{n}},$$

so by Thm. 1.5, Lem. 4.1, Lem. 4.4, and Cor. 4.5 (b), the processes  $(|X_{\mathbf{b}(\lfloor nt \rfloor)}|/(\sigma\sqrt{n}))_{t \geq 0}$  converge in law to a reflected Brownian motion. On the other hand, Cor. 4.5 (a) implies that  $n^{-1} \max_{0 \leq t \leq 1} (\mathbf{b}(\lfloor nt \rfloor) - \lfloor nt \rfloor) \rightarrow 0$  in probability, so we obtain the CLT for the processes  $(|X_{\lfloor nt \rfloor}|/(\sigma\sqrt{n}))_{t \geq 0}$  from the a.s. uniform continuity of Brownian motion on compact intervals.  $\square$

In the remainder of this subsection we prove Lem. 4.4 and Cor. 4.5. Let  $\mathcal{C}_{o,\ell} \equiv \mathcal{C}(o \leftrightarrow \ell)$  denote the conductance between  $o$  and  $D_\ell$  in  $\mathcal{T}$ , with respect to the stationary measure  $\varpi$  for  $\text{RW}_\rho(\mathcal{T})$  with the normalization  $\varpi(o) = d_o$ . We will make use of the following conductance lower bound:

**Lemma 4.6.** *Under (H1), (H2), and (H3<sup>2</sup>), there exist  $0 < r, C < \infty$  such that for all  $\epsilon > 0$ ,  $\text{MGW}(\mathcal{C}_{o,k}^{-1} \geq k^{1+\epsilon} \mid \mathbb{X}^c) \leq Ck^{-r\epsilon}$ .*

The proof of the lemma is deferred to §6.3 where we also provide a quenched conductance lower bound (see Propn. 6.5) which is not needed in the proof of the main theorem. Lem. 4.6 readily implies an upper bound on the amount of time

$$N_\alpha(n) \equiv \sum_{t=0}^n \mathbf{1}_{\{|X_t| \leq n^\alpha\}}, \quad N_\alpha^\diamond(n) \equiv \sum_{t=0}^n \mathbf{1}_{\{d(Y_t, \xi) \leq n^\alpha\}}$$

spent by  $X$  (resp.  $Y$ ) within distance  $n^\alpha$  of the root (resp. marked ray) by time  $n$ :

**Corollary 4.7.** *Assume (H1), (H2), and (H3<sup>2</sup>). Then*

$$\left. \begin{array}{l} \text{MGW}(N_\alpha(n) \geq n^{1/2+\alpha+\epsilon} \mid \mathbb{X}^c) \\ \text{IMGW}_0(N_\alpha^\diamond(n) \geq n^{1/2+\alpha+\epsilon}) \end{array} \right\} \leq Cn^{-\alpha\epsilon}.$$

*Proof.* By iterated expectations, Markov's inequality, and Lem. 3.5,

$$\begin{aligned} \text{MGW}(N_\alpha(n) \geq n^{1/2+\alpha+2\epsilon} \mid \mathbb{X}^c) &\leq n^{-\epsilon} + \text{MGW}(\mathbb{E}_{\mathcal{T}}[N_\alpha(n)] \geq n^{1/2+\alpha+\epsilon} \mid \mathbb{X}^c) \\ &\leq n^{-\epsilon} + Ce^{-n^{2\epsilon}/3} + \text{MGW}(\{\tau_{n,\epsilon}^{\text{hit}} > n\} \cap \{\mathbb{E}_{\mathcal{T}}[N_\alpha(\tau_{n,\epsilon}^{\text{hit}})] \geq n^{1/2+\alpha+\epsilon}\} \mid \mathbb{X}^c) \end{aligned}$$

where  $\tau_{n,\epsilon}^{\text{hit}} \equiv \inf\{t \geq 0 : |X_t| = \lfloor n^{1/2+\epsilon} \rfloor\}$ . Wald's identity gives

$$\mathbb{E}_{\mathcal{T}}[N_\alpha(\tau_{n,\epsilon}^{\text{hit}})] \leq \mathbb{E}_{\mathcal{T}}[L_o(\tau_{n,\epsilon}^{\text{hit}})] \mathbb{E}_{\mathcal{T}}[N_\alpha(\tau^\circ)]$$

with  $L_A(n)$  the number of visits the walk makes to set  $A$  by time  $n$ . Recalling (3.8), we have

$$\mathbb{E}_{\mathcal{T}}[L_o(\tau)] \leq \frac{d_o}{\mathcal{C}(o \leftrightarrow \lfloor n^{1/2+\epsilon} \rfloor)}, \quad \mathbb{E}_{\mathcal{T}}[N_{\tau^\circ}(\alpha)] \leq C \sum_{k=0}^{\lfloor n^\alpha \rfloor + 1} |D_k|,$$

so the MGW bound follows from Lem. 4.6 with a few more applications of Markov's inequality.

For the IMGW<sub>0</sub> bound we argue as in the proof of Lem. 3.9: by Markov's inequality and Cor. 3.6,

$$\begin{aligned} \text{IMGW}_0(N_\alpha^\diamond(n) \geq n^{1/2+\alpha+2\epsilon}) &\leq n^{-\epsilon} + \text{IMGW}_0(\mathbb{E}_{(\mathcal{T}, \xi)}[N_\alpha^\diamond(n)] \geq n^{1/2+\alpha+\epsilon}) \\ &\leq n^{-\epsilon} + Ce^{-n^{4\epsilon}/3} + \text{IMGW}_0(N_\alpha^\diamond(\tau_{n,\epsilon}^{\text{hit}}) \geq n^{1/2+\alpha+2\epsilon}), \end{aligned}$$

so it suffices to bound the last term. By Wald's identity and (3.11),

$$\begin{aligned} \mathbb{E}_{(\mathcal{T}, \xi)}[N_\alpha^\diamond(\tau_{n,\epsilon}^{\text{hit}})] &\leq \sum_{i=0}^{\lfloor n^{1/2+\epsilon} \rfloor - 1} \mathbb{E}_{(\mathcal{T}, \xi)}[L_i(\tau_{n,\epsilon}^{\text{hit}})] \mathbb{E}_{(\mathcal{T}, \xi)}^i[N_\alpha^\diamond(\tau^{\text{exc}})] \\ &\leq C \sum_{i=0}^{\lfloor n^{1/2+\epsilon} \rfloor - 1} d_i \mathbb{E}_{(\mathcal{T}, \xi)}^i[N_\alpha^\diamond(\tau^{\text{exc}})], \end{aligned}$$

so again the bound follows by using Markov's inequality.  $\square$

Most of the technical estimates required for the proof of Lem. 4.4 are contained in the following auxiliary lemma (cf. [12, Lem. 7.3]). For  $\ell(n)$  as in (4.1), define the sequence of  $(\mathcal{H}_s)$ -stopping times

$$\Theta_0 \equiv 0, \quad \Theta_{j+1} \equiv \min\{s > \Theta_j : ||X_s^{\text{int}}| - |X_{\Theta_j}^{\text{int}}|| = \ell(n)\}$$

and similarly the sequence of  $(\mathcal{H}_s^\diamond)$ -stopping times

$$\Theta_0^\diamond \equiv 0, \quad \Theta_{j+1}^\diamond \equiv \min\{s > \Theta_j^\diamond : |d(Y_s^{\text{int}}, \xi) - d(Y_{\Theta_j^\diamond}^{\text{int}}, \xi)| = \ell(n)\}.$$

**Lemma 4.8.** *Assume (H1) and (H2).*

(a) *Assume (H3<sup>p</sup>) and let*

$$\mathbf{C} \equiv \mathbf{C}(n, \epsilon) \equiv \{v \in \mathcal{T} : W_v > n^{1/4-\epsilon}\}$$

*(well-defined for trees with and without ray). Then*

$$\left. \begin{array}{l} \text{MGW}(\tau^{\text{hit}}(\mathbf{C}) \leq n) \\ \text{IMGW}_0(\tau^{\text{hit}}(\mathbf{C} \setminus \xi) \leq n) \end{array} \right\} \leq C(p, \epsilon) n^{1-p(1/4-\epsilon)}.$$

*For  $p > 4$  the right-hand side can be made  $\leq n^{-c}$  for  $c \equiv c(p, \epsilon) > 0$ .*

(b) *Assuming (H3<sup>p</sup>) with  $p > 2$ , for any  $\epsilon > 0$  there exists  $c \equiv c(p, \epsilon)$  such that*

$$\left. \begin{array}{l} \text{MGW}(I_n \geq n^{1/2+\epsilon} \mid \mathbb{X}^c) \\ \text{IMGW}_0(I_n^\diamond \geq n^{1/2+\epsilon}) \end{array} \right\} \leq e^{-cn^{\epsilon/2}}.$$

(c) With  $\Theta_j, \Theta_j^\diamond$  defined as above,

$$\left. \begin{array}{l} \text{MGW}(\mathbf{t}(\Theta_{3I_n}) \leq n) \\ \text{IMGW}_0(\mathbf{t}^\diamond(\Theta_{3I_n}^\diamond) \leq n) \end{array} \right\} \leq e^{-c\ell(n)}.$$

(d) Recalling the notation of (3.7) and (3.9), define

$$\mathbf{A} \equiv \mathbf{A}(n, \alpha, \epsilon) \equiv \bigcup_{k=0}^{\lfloor (\log n)^2 \rfloor} A_{\lfloor n^\alpha \rfloor - k}^\epsilon, \quad \mathbf{B} \equiv \mathbf{B}(n, \alpha, \epsilon) \equiv \bigcup_{k=0}^{\lfloor (\log n)^2 \rfloor} B_{\lfloor n^\alpha \rfloor - k}^\epsilon.$$

Assuming  $(\text{H}3^p)$  with  $p > 2$ , there exists  $\alpha_0(p) \in (0, 1/2)$  such that for all  $\alpha \geq \alpha_0(p)$  and all  $\epsilon < \epsilon_0(p, \alpha)$  (with  $\epsilon_0(p, \alpha) > 0$ ),

$$\left. \begin{array}{l} \text{MGW}(\tau_{\mathbf{A}}^{\text{hit}} \leq n) \\ \text{IMGW}_0(\tau_{\mathbf{B}}^{\text{hit}} \leq n) \end{array} \right\} \leq n^{-c} \quad (4.3)$$

for some  $c \equiv c(p, \alpha, \epsilon) > 0$ . If further  $p > 4$  then  $\alpha_0, \epsilon_0$  can be chosen such that

$$\text{MGW}(\tau_{\mathbf{A}_+}^{\text{hit}} \leq n) \leq n^{-c}, \quad \mathbf{A}_+ \equiv \mathbf{A}_+(n, \alpha, \epsilon) \equiv \bigcup_{k \geq \lfloor n^\alpha \rfloor} A_k^\epsilon. \quad (4.4)$$

*Proof.* (a) See proof of [31, (63)].

(b) We will show that with probability  $\geq 1 - e^{-cn^{\epsilon/2}}$  conditioned on  $\mathbb{X}^c$ , one of the first  $\lfloor n^{1/2+\epsilon}/2 \rfloor$  excursions has length  $\eta_i - \tau_i > n$  which certainly implies the result. Conditioning on  $\mathbb{X}^c$  is needed simply to ensure  $\tau^\mathbb{X} = \infty$ ; for the purpose of proving the claim we may artificially define  $\eta_i - \tau_i = \infty$  for  $i > \tau^\mathbb{X}$ .

Then, conditioned on  $(\eta_j - \tau_j)_{j=1}^{i-1}$ , the probability that  $\eta_i - \tau_i > n$  is bounded below by a constant times  $\text{MGW}(\tau^\diamond > n)$ . Further

$$\mathbb{P}_{\mathcal{T}}(\tau^\diamond > n) \geq \mathbb{P}_{\mathcal{T}}(\tau^\diamond > \tau_{n, \epsilon/2}^{\text{hit}} > n) \geq \mathbb{P}_{\mathcal{T}}(\tau^\diamond > \tau_{n, \epsilon/2}^{\text{hit}}) - \mathbb{P}_{\mathcal{T}}(\tau_{n, \epsilon/2}^{\text{hit}} \leq n),$$

so Lem. 3.5 and Lem. 4.6 imply  $\text{MGW}(\tau^\diamond > n) \geq c/n^{1/2+\epsilon/2}$  for  $c \equiv c(p, \epsilon)$ . Thus the probability that none of the first  $\lfloor n^{1/2+\epsilon} \rfloor$  excursions has length  $> n$  is

$$\leq \left(1 - \frac{c}{n^{1/2+\epsilon/2}}\right)^{\lfloor n^{1/2+\epsilon}/2 \rfloor} \leq e^{-cn^{\epsilon/2}},$$

which proves the result.

(c) We follow the proof of [31, Lem. 11]. On the  $\text{IMGW}_0$  tree, since  $d(Y_0, \xi) = 0$ ,  $d(Y^{\text{int}}, \xi)$  must increase by  $\ell(n)$  going from  $\Theta_{j-1}^\diamond$  to  $\Theta_j^\diamond$  for at least half of the indices  $j \leq 3I_n^\diamond$  so  $\{\mathbf{t}(\Theta_{3I_n}) \leq n\}$  implies the event

$$\mathbf{G}^\diamond \equiv \{\exists i \leq I_n^\diamond, \Theta_{j-1}^\diamond, \Theta_j^\diamond \in J_j^\diamond, d(Y_{\Theta_j^\diamond}^{\text{int}}, \xi) > d(Y_{\Theta_{j-1}^\diamond}^{\text{int}}, \xi)\}.$$

This in turn implies one of two possibilities:

1. there exist times  $t_0 < t_1 < t_2 \leq n$  with  $Y_{t_0} = Y_{t_2}$  and  $d(Y_{t_0}, \xi) = d(Y_{t_1}, \xi) + \ell(n)/4$ , or
2. there exist times  $t_1 < t_2 \leq n$  with  $d(Y_{t_2}, \xi) = d(Y_{t_1}, \xi) + \ell(n)/4$  such that  $a_0$  does not appear on the geodesic between the  $Y_{t_i}$ .

By a random walk estimate (cf. (3.8)) summed over at most  $n^2$  possibilities for  $(Y_{t_0}, Y_{t_1})$ , the first event has probability  $\leq Cn^2\rho^{-\ell(n)/4}$ . The second event has probability  $\leq e^{-c\ell(n)}$  by the construction of  $\text{IMGW}_0$  and the irreducibility of the Markov chain, and combining these estimates gives the bound for  $\text{IMGW}_0$ . The bound for  $\text{MGW}$  follows by a similar argument.

(d) We first prove the bounds for  $\text{MGW}$ ; the argument is similar to that of Cor. 4.7: again it suffices to bound  $\text{MGW}(\tau_{\mathbf{A}}^{\text{hit}} \leq \tau_{n,\epsilon}^{\text{hit}})$ , and Wald's identity gives

$$\mathbb{P}_{\mathcal{T}}(\tau_{\mathbf{A}}^{\text{hit}} \leq \tau_{n,\epsilon}^{\text{hit}}) \leq \mathbb{E}_{\mathcal{T}}[L_o(\tau_{n,\epsilon}^{\text{hit}})] \mathbb{P}_{\mathcal{T}}[\tau_{\mathbf{A}}^{\text{hit}} < \tau^{\circ}].$$

By (3.8) and Markov's inequality,  $\mathbb{E}_{\mathcal{T}}[L_o(\tau_{n,\epsilon}^{\text{hit}})] \leq n^{1/2+2\epsilon}$  except with probability at most  $n^{-c}$  for  $c \equiv c(p, \epsilon) > 0$ . By Lem. 3.8,

$$\text{MGW}(\tau_{\mathbf{A}}^{\text{hit}} < \tau^{\circ}) \leq C(\alpha, p, \epsilon)(\log n)^2(n^{\alpha})^{-p/2}.$$

Since  $p > 2$ , we can choose  $\alpha$  sufficiently close to  $1/2$  and  $\epsilon$  sufficiently small such that Markov's inequality gives  $\text{MGW}(\mathbb{P}_{\mathcal{T}}(\tau_{\mathbf{A}}^{\text{hit}} < \tau^{\circ}) \geq n^{-(1/2+3\epsilon)}) \leq n^{-c}$ , from which (4.3) follows for  $\text{MGW}$ . (4.4) follows similarly by noting

$$\text{MGW}(\tau_{\mathbf{A}_+}^{\text{hit}} < \tau^{\circ}) \leq C(\alpha, p, \epsilon)(n^{\alpha})^{1-p/2}.$$

The bound (4.3) for  $\text{MGW}$  together with the argument of Cor. 4.7 gives

$$\text{IMGW}_0(\tau_{\mathbf{B}}^{\text{hit}} \leq \tau_{n,\epsilon}^{\text{hit}}) \leq C(p, \epsilon)(\log n)^2 n^{1/2+\epsilon-\alpha p/2},$$

and the bound (4.3) for  $\text{IMGW}_0$  follows by choosing  $\alpha$  close to  $1/2$  and  $\epsilon$  small.  $\square$

*Proof of Lem. 4.4.* We modify the proof of [31, (55)] (see also [12, (28)]; our Lem. 4.8 plays the role of [12, Lem. 5.1]).

On the  $\text{MGW}$  tree write  $\mathbf{E} \equiv \{\Delta_n(\alpha) \neq \Delta_n\} = \{\max_{s \leq \Delta_n} |X_s^{\text{int}}| \geq n^{\alpha}\}$ . If we define

$$\Upsilon \equiv \{I_n \geq n^{1/2+\epsilon}\} \cup \{\mathbf{t}(\Theta_{3I_n} \leq n)\} \cup \{\tau_{\mathbf{A}}^{\text{hit}} \leq n\} \cup \{\tau_{\mathbf{C}}^{\text{hit}} \leq n\}$$

and consider the process  $\mathcal{M}(j) \equiv S_{X^{\text{int}}(\Theta_j)}$ , we have

$$\text{MGW}(\mathbf{E}) \leq \text{MGW}(\Upsilon) + \text{MGW}\left(\Upsilon^c \cap \left\{ \max_{j \leq 3n^{1/2+\epsilon}} \mathcal{M}(j) \geq (\eta - \epsilon)(n^{\alpha} - \ell(n)) \right\}\right). \quad (4.5)$$

By Lem. 4.8 it suffices to bound the second term. To this end let  $\tau_i^{\circ}$  denote the  $i$ -th return of  $(X_{\Theta_j}^{\text{int}})_j$  to the root (with  $\tau_0^{\circ} \equiv 0$ ): then for each  $i \geq 0$  the process  $\mathcal{M}_i(j) \equiv \mathcal{M}(j \wedge \tau_{i+1}^{\circ})$  is a supermartingale for  $j \geq \tau_i^{\circ} + 1$ , and the second term of (4.5) is

$$\leq \text{MGW}\left(\Upsilon^c \cap \left\{ \max_{i: \tau_i^{\circ} \leq 3n^{1/2+\epsilon}} \max_{j \leq 3n^{1/2+\epsilon}} [\mathcal{M}_i(\tau_i^{\circ} + j) - \mathcal{M}_i(\tau_i^{\circ} + 1)] \geq \frac{\eta n^{\alpha}}{2} \right\}\right)$$

(for  $n$  large and suitable  $\alpha, \epsilon$ ). If we define the  $(\mathcal{H}_{\Theta_j})_j$ -stopping time

$$\Psi \equiv \inf\{j : \mathbf{t}(\Theta_j) > \tau_{\mathbf{C}}^{\text{hit}}\},$$

then

$$\mathfrak{M}_i(j) \equiv \mathcal{M}_i(j \wedge \Psi) - [\mathcal{M}_i(\Psi) - \mathcal{M}_i(\Psi - 1)]\mathbf{1}_{\{\Psi \leq j\}}, \quad j \geq \tau_i^{\circ} + 1$$



is a supermartingale with differences  $\leq \ell(n)n^{1/4-\epsilon}$ , so the Azuma–Hoeffding inequality gives

$$\mathbb{P}_{\mathcal{T}}\left(\max_{j \leq 3n^{1/2+\epsilon}} [\mathfrak{M}_i(\tau_i^\circ + j) - \mathfrak{M}_i(\tau_i^\circ + 1)] \geq \frac{\eta n^\alpha}{2}\right) \leq \exp\left\{-\frac{(\eta n^\alpha/2)^2}{2(3n^{1/2+\epsilon})\ell(n)^2 n^{2(1/4-\epsilon)}}\right\}.$$

Choosing  $\alpha, \epsilon$  appropriately and summing over at most  $3n^{1/2+\epsilon}$  return times  $\tau_i^\circ$  gives the desired bound on the second term of (4.5), from which the **MGW** bound follows.

The bound for the **IMGW**<sub>0</sub>-probability of  $\mathbf{E}^\circ \equiv \{\Delta_n^\circ(\alpha) \neq \Delta_n^\circ\}$  is similar, indeed simpler since  $\mathcal{M}^\circ(j) \equiv S_{Y^{\text{int}}(\Theta_j^\circ)}$  is always a supermartingale. With

$$\Upsilon^\circ \equiv \{I_n^\circ \geq n^{1/2+\epsilon}\} \cup \{\mathbf{t}^\circ(\Theta_{3I_n^\circ}^\circ) \leq n\} \cup \{\tau_{\mathbf{B}}^{\text{hit}} \leq n\} \cup \{\tau_{\mathbf{C}}^{\text{hit}} \leq n\},$$

we have from Lem. 4.8 that **IMGW**<sub>0</sub>( $\mathbf{E}^\circ$ ) is

$$\leq n^{-c} + \text{IMGW}_0\left((\Upsilon^\circ)^c \cap \left\{\max_{j \leq 3n^{1/2+\epsilon}} \mathcal{M}^\circ(j) \geq \frac{\eta n^\alpha}{2}\right\}\right),$$

and applying the Azuma–Hoeffding bound gives the result.  $\square$

*Proof of Cor. 4.5. (a)* We have the set inclusions

$$\begin{aligned} \{\Delta_n \geq n^{1/2+\alpha+\epsilon}\} &\subseteq \{\Delta_n \neq \Delta_n(\alpha)\} \cup \{N_\alpha(n) \geq n^{1/2+\alpha+\epsilon}\}, \\ \{\Delta_n^\circ \leq n^{1/2+\alpha+\epsilon}\} &\subseteq \{\Delta_n^\circ \neq \Delta_n^\circ(\alpha)\} \cup \{N_\alpha^\circ(n) \geq n^{1/2+\alpha+\epsilon}\}, \end{aligned}$$

so the result follows from Lem. 4.4 and Cor. 4.7.

(b) Let  $(h_s)_{s \geq 0}$  denote the height process for the walk  $Y$  restricted to  $\xi$ , i.e. erasing all excursions away from  $\xi$ ; clearly  $\mathbb{D}_n \leq \mathbb{D}'_n \equiv \max\{h_s - h_r : 0 \leq r \leq s \leq n\}$ . But  $h_s$  is simply a random walk on  $\mathbb{Z}_{\leq 0}$  with a  $\rho$ -bias in the negative direction. Set  $\sigma_0 \equiv 0$ ,

$$\sigma_j \equiv \inf\{s > \sigma_{j-1} : h_s = h_{\sigma_{j-1}} - 1\}, \quad j \geq 1.$$

Now the processes  $(\bar{h}_s^{(j)} \equiv h_s - h_{\sigma_j})_{\sigma_j \leq s \leq \sigma_{j+1}}$  are i.i.d., and clearly  $\sigma_n \geq n$ , so

$$\mathbb{D}'_n \leq \max_{0 \leq j < n} \left(\max_s \bar{h}_s^{(j)}\right).$$

The probability of  $\max_s \bar{h}_s^{(j)} \geq m$  is at most the probability that a random walk on  $\mathbb{Z}$  started at 0 with a  $\rho$ -bias in the negative direction will reach  $m$  before  $-1$ , which is  $(1 - \rho^{-1})/(\rho^m - \rho^{-1}) \leq \rho^{-m}$ . Summing over  $j$  gives  $\mathbb{P}_{(\mathcal{T}, \xi)}(\mathbb{D}'_n \geq m) \leq n\rho^{-m}$ , **IMGW**<sub>0</sub>-a.s.  $\square$

**4.3. Quenched MGW-CLT.** We now describe how to move from the annealed to the quenched CLT; the proof is motivated by ideas in [31, §6-7] and [5, Lem. 4.1]. For given  $n \geq 1$ , let  $\mathbf{s}$  denote the unique increasing bijection

$$\mathbf{s} : \mathbb{Z}_{\geq 0} \rightarrow \bigcup_{i \geq 1} [\tau_i, \eta_i),$$

and let  $X_t^{\text{exc}} \equiv X_{\mathbf{s}(t)}$ , the excursion process of  $X$  with parameter  $n$  (recalling Rmk. 4.2). For  $\mathbf{s}(t) \in [\tau_i, \eta_i)$  write  $X_t^{\text{cent}} \equiv |X_t^{\text{exc}}| - |X_{\tau_i}|$ .

*Proof of Thm. 1.1.* We show the quenched CLT for  $X$  through a quenched CLT for  $X^{\text{cent}}$  along geometrically increasing subsequences  $b_k \equiv \lfloor b^k \rfloor$  ( $k \geq 0$ ) with  $b > 1$ .

*Step 1: annealed CLT for  $X^{\text{cent}}$ .*

The time killed during the first  $n$  steps of  $X$  is  $n - \mathbf{s}^{-1}(n) \leq \Delta_n$ , so Cor. 4.5 (a) gives  $n^{-1} \sup_{0 \leq t \leq T} |\mathbf{s}(\lfloor nt \rfloor) - \lfloor nt \rfloor| \rightarrow 0$  in MGW-probability. It follows from Propn. 4.3 and the continuity of Brownian motion that the processes  $X_{\lfloor nt \rfloor}^{\text{cent}}/(\sigma\sqrt{n})$  also satisfy the annealed MGW-CLT.

*Step 2: quenched CLT for  $X^{\text{cent}}$  along geometrically increasing subsequences.*

Recalling Rmk. 1.2, let  $\mathbb{B}^n(X) \equiv (\mathbb{B}_t^n(X))_{t \geq 0}$  denote the polygonal interpolation of  $j/n \mapsto X_j^{\text{cent}}/(\sigma\sqrt{n})$ , and regard  $\mathbb{B}^n(X)$  as an element of  $C[0, T]$  with the norm

$$d_T(u, u') \equiv \left( \sup_{0 \leq t \leq T} |u_t - u'_t| \right) \wedge 1.$$

We will show that for all Lipschitz functions  $F : C[0, T] \rightarrow [-1, 1]$  with Lipschitz constant  $\leq 1$ ,

$$\sum_{k \geq 0} \text{Var}_{\text{MGW}}[\mathbb{E}_{\mathcal{T}}[F[\mathbb{B}^{\lfloor b^k \rfloor}(X)]]] < \infty. \quad (4.6)$$

The Borel–Cantelli lemma then implies (cf. [5, Lem. 4.1]) that for MGW-a.e.  $\mathcal{T}$ , the processes  $X_{\lfloor nt \rfloor}^{\text{cent}}/(\sigma\sqrt{n})$  converge in law to the absolute value of a standard Brownian motion along the subsequence  $b_k$ .

To see (4.6), let  $\mathcal{T} \sim \text{MGW}$ , let  $(X^i, \mathbf{s}^i)$  be two independent realizations of  $(X, \mathbf{s})$  conditioned on  $\mathcal{T}$ , and write  $\mathbb{B}^{n,i} \equiv \mathbb{B}^n(X^i)$ . Then

$$\text{Var}_{\text{MGW}}[\mathbb{E}_{\mathcal{T}}[F[\mathbb{B}^n(X)]]] = \mathbb{E}_{\text{MGW}}[F(\mathbb{B}^{n,1})F(\mathbb{B}^{n,2})] - \mathbb{E}_{\text{MGW}}[F(\mathbb{B}^{n,1})]^2.$$

Let  $\mathcal{E}_n^i$  denote the subtree explored by  $X^i$  up to time  $n$ . Conditioning on the first  $\ell(n)/2$  levels of  $\mathcal{T}$ , let  $(\dot{\mathcal{E}}_n^i, \dot{X}_n^i, \dot{\mathbf{s}}^i)$  ( $i = 1, 2$ ) be two independent realizations of  $(\mathcal{E}_n^i, X^i, \mathbf{s}^i)$ : then the processes  $(\dot{X}_n^i)^{\text{cent}}$  are *exactly* independent with law not depending on the first  $\ell(n)/2$  levels of  $\mathcal{T}$ . Moreover, if  $\mathcal{A}_n$  denotes the event that the paths of  $\dot{X}^1$  and  $\dot{X}^2$  up to time  $\max_i (\dot{\mathbf{s}}^i)^{-1}(n)$  have no common vertices at distance more than  $\ell(n)/2$  from the root, then we can couple  $(\mathcal{E}_n^i, X^i|_{[0,n]})_{i=1,2}$  with  $(\dot{\mathcal{E}}_n^i, \dot{X}_n^i|_{[0,n]})_{i=1,2}$  such that the processes agree on the event  $\mathcal{A}_n$ . Therefore

$$\begin{aligned} \text{Var}_{\text{MGW}}[\mathbb{E}_{\mathcal{T}}[F[\mathbb{B}^n(X)]]] &\leq \mathbb{E}_{\text{MGW}}[F(\dot{\mathbb{B}}^{n,1})F(\dot{\mathbb{B}}^{n,2})] + \text{MGW}(\mathcal{A}_n) - \mathbb{E}_{\text{MGW}}[F(\mathbb{B}^{n,1})]^2 \\ &= \mathbb{E}_{\text{MGW}}[F(\dot{\mathbb{B}}^{n,1})]^2 + \text{MGW}(\mathcal{A}_n) - \mathbb{E}_{\text{MGW}}[F(\mathbb{B}^{n,1})]^2 = \text{MGW}(\mathcal{A}_n) \end{aligned}$$

We claim  $\text{MGW}(\mathcal{A}_n) \leq n^{-c}$ : since Cor. 4.7 and Cor. 4.5 (a) imply  $\text{MGW}(2n - \mathbf{s}^{-1}(2n) \geq n) \leq n^{-c}$ , it suffices to bound the probability that the paths of  $X^1$  and  $X^2$  up to time  $2n$  intersect at distance  $> \ell(n)/2$  from the root. But the chance that  $X^2$  hits a given vertex  $v$  with  $|v| > \ell(n)/2$  by time  $2n$  is  $\leq Cn\rho^{-\ell(n)/2}$ , and summing over the vertices visited by  $X^1$  proves the claim. The variance condition (4.6) now follows by summing over  $(b_k)_{k \geq 0}$ .

*Step 3: quenched CLT for  $X$  along geometrically increasing subsequences.* Extend  $\mathbf{s}^{-1}$  to a nondecreasing map  $\mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$  by setting  $\mathbf{s}^{-1}(t) = \mathbf{s}^{-1}(\tau_i)$  for  $t \in [\eta_{i-1}, \tau_i)$ :

then

$$||X_t| - X_{\mathbf{s}^{-1}(t)}^{\text{cent}}| = \begin{cases} |X_{\tau_i}|, & t \in [\tau_i, \eta_i), \\ |X_t|, & t \in [\eta_{i-1}, \tau_i). \end{cases}$$

It follows from Lem. 4.4 and Cor. 4.5 (a) that for any  $b > 1$ ,

$$\left. \begin{aligned} & b_k^{-1} \sup_{0 \leq t \leq b_k T} \{ \mathbf{s}^{-1}(t) - \mathbf{s}^{-1}(t) \} \\ & b_k^{-1/2} \sup_{0 \leq t \leq b_k T} ||X_t| - |X_{\mathbf{s}^{-1}(t)}^{\text{cent}}|| \end{aligned} \right\} \xrightarrow{k \rightarrow \infty} 0, \quad \text{MGW-a.s.}$$

It follows that the processes  $|X_{nt}|/(\sigma\sqrt{n})$  satisfy the quenched CLT along the subsequence  $(b_k)_{k \geq 0}$  for any  $b > 1$ .

*Step 4: quenched CLT for  $X$  along full sequence.* For the processes  $(|X_{[nt]}|/(\sigma\sqrt{n}))_{t \geq 0}$ , MGW-a.s. tightness and convergence of finite-dimensional distributions both follow from the scaling relation

$$\mathbb{B}_t^n(X) = \sqrt{\frac{b_k}{n}} \mathbb{B}_{b_k}^{b_k} \left( t \frac{n}{b_k} \right),$$

(cf. proof of [5, Lem. 4.1]). □

*Proof of Cor. 1.3.* Given  $(\mathcal{T}, X) \sim \text{MGW} \otimes \text{RW}_\rho$  we can obtain  $(\mathcal{T}, X^{\text{cts}}) \sim \text{MGW} \otimes \text{RW}_\rho^{\text{cts}}$  by taking  $(E_i)_{i \geq 1}$  i.i.d. exponential random variables with unit mean independent of  $X$ , and setting

$$X_t^{\text{cts}} = X_{\boldsymbol{\theta}(t)}, \quad \boldsymbol{\theta}(t) = \max \left\{ i : \sum_{j=1}^i \frac{E_j}{\rho + d_{X_{j-1}}} \leq t \right\};$$

similarly we can obtain  $((\mathcal{T}^\diamond, \xi), Y^{\text{cts}}) \sim \text{IMGW}_0 \otimes \text{RW}_\rho^{\text{cts}}$  from  $((\mathcal{T}^\diamond, \xi), Y) \sim \text{IMGW}_0 \otimes \text{RW}_\rho$ . Thus a shifted coupling of  $(\mathcal{T}, X)$  with  $((\mathcal{T}^\diamond, \xi), Y)$  (as constructed in §4) naturally gives rise to a shifted coupling of  $(\mathcal{T}, X^{\text{cts}}) \sim \text{MGW} \otimes \text{RW}_\rho^{\text{cts}}$  with  $((\mathcal{T}^\diamond, \xi), Y^{\text{cts}}) \sim \text{IMGW}_0 \otimes \text{RW}_\rho^{\text{cts}}$  by using sequences  $(E_i)_{i \geq 1}$  for  $X^{\text{cts}}$  and  $(E_i^\diamond)_{i \geq 1}$  for  $Y^{\text{cts}}$  which are marginally i.i.d. exponential but such that the jump times match during the coupled excursions.

By Thm. 1.4 (b) and the exponential decay of the  $E_i^\diamond$ , it holds IMGWR-a.s. that

$$\frac{1}{n} \sum_{i=1}^n \frac{E_i^\diamond}{\rho + d_{Y_i}} \rightarrow \mathbb{E}_{\text{IMGWR}} \left[ \frac{1}{\rho + d_o} \right] = \frac{1}{2\rho}.$$

From this it is easy to see that  $n^{-1} \sup_{0 \leq t \leq T} [\boldsymbol{\theta}(nt) - 2\rho nt] \rightarrow 0$  IMGWR-a.s., so on IMGWR-a.e.  $(\mathcal{T}, \xi)$  the processes  $(h(Y_{[nt]}^{\text{cts}})/(\sigma\sqrt{2\rho n}))_{t \geq 0}$  converge in law to standard Brownian motion. The quenched MGW-CLT for  $X^{\text{cts}}$  follows from the proof of Thm. 1.1. □

5. TRANSIENCE-RECURRENCE BOUNDARY FOR  $\text{RWRE}_\lambda$ 

We now prove Thm. 1.6. Our proof is a straightforward adaptation of that of [25, Thm. 1] or [12, Propn. 1.1] once we supply the needed large deviations estimate (Lem. 5.2) on the conductances at the  $n$ -th level of the tree, extending the estimates of [25, p. 129] and [12, p. 7] to our setting of Markovian dependency.

Let  $\mathcal{D} \equiv \{\gamma : \max_{a,b} \bar{A}^{(\gamma)}(a,b) < \infty\}$ , where  $\bar{A}^{(\gamma)}$  is as defined in (1.3). Recall that  $\bar{\rho}(\gamma)$  denotes the Perron–Frobenius eigenvalue of  $\bar{A}^{(\gamma)}$  with  $\bar{\rho}(\gamma) \equiv \infty$  for  $\gamma \notin \mathcal{D}$ . The following lemma collects some basic properties of  $\bar{\rho}$ .

**Lemma 5.1.** *Under the hypotheses of Thm. 1.6,  $\bar{\rho}$  is lower semi-continuous and log-convex on  $\mathbb{R}$ , and differentiable on  $\mathcal{D}$ .*

*Proof.* Lower semi-continuity of  $\bar{\rho}$  in  $\mathcal{D}$  follows from Fatou’s lemma, and lower semi-continuity outside the closure of  $\mathcal{D}$  is trivial, so it remains to consider the boundary of  $\mathcal{D}$ : we must show that if  $\gamma \rightarrow \gamma_\infty$  with  $\max_{a,b} \bar{A}^{(\gamma)}(a,b) \rightarrow \infty$  then  $\bar{\rho}(\gamma) \rightarrow \infty$ . Recall the min-max characterization (see e.g. [15, Cor. 8.3.3])

$$\bar{\rho}(\gamma) = \max_{x \geq 0, x \neq 0} \min_{a: x_a \neq 0} \frac{(\bar{A}^{(\gamma)} \underline{x})_a}{x_a}.$$

Since  $\rho[(A^{(\gamma)})^k] = \rho(A^{(\gamma)})^k$  ( $k \in \mathbb{N}$ ), and  $\bar{A}^{(0)}$  is positive regular which implies  $\bar{A}^{(\gamma)}$  is also for all  $\gamma \in \mathcal{D}$ , we may assume without loss that  $\min_{a,b} \bar{A}^{(\gamma)}(a,b) \geq \epsilon$  for all  $\gamma$  in a neighborhood of  $\gamma_\infty$ . Applying the min-max characterization to the vectors  $\underline{x} = \mathbf{1}_a$  gives  $\bar{\rho}(\gamma) \geq \max_a \bar{A}^{(\gamma)}(a,a)$ . Applying it to the vectors

$$\underline{x} = \left( \mathbf{1}_{\{c=a\}} + \mathbf{1}_{\{c=b\}} \frac{\sqrt{\epsilon}}{\sqrt{A^{(\gamma)}(a,b)}} \right)_{c \in \Omega}, \quad a \neq b$$

gives

$$\begin{aligned} \bar{\rho}(\gamma) &\geq \left( A^{(\gamma)}(a,a) + \sqrt{\epsilon A^{(\gamma)}(a,b)} \right) \wedge \left( A^{(\gamma)}(b,a) \frac{\sqrt{A^{(\gamma)}(a,b)}}{\sqrt{\epsilon}} + A^{(\gamma)}(b,b) \right) \\ &\geq \sqrt{\epsilon A^{(\gamma)}(a,b)}. \end{aligned}$$

Combining gives  $\bar{\rho}(\gamma)^2 \geq \epsilon \max_{a,b \in \Omega} A^{(\gamma)}(a,b)$  which proves lower semi-continuity.

The entries of  $\bar{A}^{(\gamma)}$  are log-convex in  $\gamma$  by Hölder’s inequality, so  $\bar{\rho}$  is log-convex by monotonicity and log-convexity of the Perron–Frobenius eigenvalue in the entries of the matrix (see e.g. [15, Cor. 8.1.19] and [6, Exercise 4.34]). For differentiability of  $\bar{\rho}$  in  $\mathcal{D}$  see [9, p. 75].  $\square$

For  $\gamma \in \mathcal{D}$  let  $\underline{e}^{(\gamma)}$  and  $\underline{g}^{(\gamma)}$  denote the associated left and right Perron–Frobenius eigenvectors; we use the shorthand

$$\underline{e} \equiv \underline{e}^{(0)}, \quad \underline{g} \equiv \underline{g}^{(0)}.$$

For  $\mathcal{T} \sim \overline{\text{MGW}}$  and  $v \in \mathcal{T}$ , let

$$C_v \equiv \prod_{o < u \leq v} \alpha_u,$$

the conductance of the edge leading to  $v$ . The natural generalization of the martingale introduced in §2.3 is

$$\mathfrak{Z}_n^{(\gamma)} = \frac{1}{\bar{\rho}(\gamma)^n} \sum_{v \in D_n} \bar{e}_{\chi_v}^{(\gamma)} C_v^\gamma; \quad (5.1)$$

this is a multi-type Mandelbrot's martingale and has been studied in various contexts, for example as the Laplace transform of the branching random walk with increments  $\log \alpha_v$  [7, 22]. Using this martingale we can make a change of measure and control the conductances at the  $n$ -th level by controlling the conductance of the edge leading to a *random* vertex: recalling (2.7), for each  $a \in \mathcal{Q}$  define the size-biased measure  $\bar{\mathcal{Q}}_n^a$  on  $\Omega$  by

$$\frac{d\bar{\mathcal{Q}}_n^a}{d\text{MGW}^a} = \frac{\mathfrak{Z}_n^{(0)}}{e_a}.$$

We then let  $\bar{\mathcal{Q}}_{n\star}^a$  denote the measure on pairs  $(\mathcal{T}, v_n)$  obtained by letting  $\mathcal{T} \sim \bar{\mathcal{Q}}_n^a$  and choosing  $v_n \in D_n$  according to weights  $e_{\chi_v}$ .

**Lemma 5.2.** *Under the hypotheses of Thm. 1.6, for each  $a \in \mathcal{Q}$ , under  $\bar{\mathcal{Q}}_{n\star}^a$  the random variables  $n^{-1} \log C_{v_n}$  satisfy a large deviation principle with good rate function  $\Lambda^*(x) \equiv \sup_{\gamma} (\gamma x - \Lambda(\gamma))$ , where  $\Lambda(\gamma) \equiv \log \bar{\rho}(\gamma) - \log \bar{\rho}(0)$ . In particular, for any  $0 < z < y$ ,*

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \bar{\mathcal{Q}}_{n\star}^a(C_{v_n} > z^n) \geq - \sup_{\gamma \geq 0} (\gamma \log y - \Lambda(\gamma)). \quad (5.2)$$

*Proof.* Fixing  $a \in \mathcal{Q}$ , let  $\Lambda_n \equiv \Lambda_n^a$  denote the cumulant generating function of  $n^{-1} \log C_{v_n}$  with respect to  $\bar{\mathcal{Q}}_{n\star}^a$ , that is,

$$\Lambda_n(\gamma) = \log \mathbb{E}_{\bar{\mathcal{Q}}_{n\star}^a} [C_{v_n}^{\gamma/n}].$$

Then

$$\begin{aligned} e^{\Lambda_n(n\gamma)} &= \mathbb{E}_{\bar{\mathcal{Q}}_{n\star}^a} [C_{v_n}^\gamma] = \mathbb{E}_{\bar{\mathcal{Q}}_n^a} \left[ \frac{\sum_{v \in D_n} e_{\chi_v} C_v^\gamma}{\sum_{v \in D_n} e_{\chi_v}} \right] = \frac{1}{e_a \bar{\rho}(0)^n} \mathbb{E}_{\text{MGW}^a} \left[ \sum_{v \in D_n} e_{\chi_v} C_v^\gamma \right] \\ &\asymp \frac{\bar{\rho}(\gamma)^n}{\bar{\rho}(0)^n} \mathbb{E}_{\text{MGW}^a} [\mathfrak{Z}_n^{(\gamma)}] \asymp \frac{\bar{\rho}(\gamma)^n}{\bar{\rho}(0)^n}, \end{aligned}$$

where  $\asymp$  indicates equivalence up to constant factors depending only on  $\underline{e}$  and  $\bar{e}^{(\gamma)}$ . Thus

$$\lim_{n \rightarrow \infty} \frac{1}{n} \Lambda_n(n\gamma) = \log \bar{\rho}(\gamma) - \log \bar{\rho}(0) = \Lambda(\gamma).$$

By Lem. 5.1 this is an essentially smooth convex function in the sense of [9, Defn. 2.3.5], so the large deviation principle follows from the Gärtner-Ellis theorem (see [9, Thm. 2.3.6]). In particular, for any  $0 < z < y$ , [9, (2.3.8)] implies

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \bar{\mathcal{Q}}_{n\star}^a(C_{v_n} > z^n) \geq - \inf_{x > \log z} \Lambda^*(x) \geq -\Lambda^*(\log y)$$

(making use of [9, Lem. 2.3.9]). The result (5.2) follows immediately if  $\log y = \Lambda'(\gamma)$  for some  $\gamma \geq 0$ , or if  $\log y \geq \sup_{\gamma \geq 0} \Lambda'(\gamma)$  in which case  $\sup_{\gamma} (\gamma \log y - \Lambda(\gamma)) =$

$\lim_{\gamma \rightarrow \infty} (\gamma \log y - \Lambda(y))$ . Next, the assumption that  $\Lambda < \infty$  in a neighborhood of 0 implies, via the relation  $\Lambda(\gamma) = \sup_x (x\gamma - \Lambda^*(x))$ , that  $\lim_{|x| \rightarrow \infty} \Lambda^*(x) = \infty$ , therefore  $\Lambda^*$  attains its global infimum at  $x_0 = \Lambda'(0)$  with  $\Lambda^*(x_0) = -\Lambda(0)$ . Therefore (5.2) again holds in the remaining case  $\log y \leq \inf_{\gamma \geq 0} \Lambda'(\gamma) = x_0$ .  $\square$

Thm. 1.6 now follows by adapting the proof of [25, Thm. 1]:

*Proof of Thm. 1.6.* Since the bias  $\lambda$  can always be absorbed into the environment variables  $\alpha_v$  ( $v \in \mathcal{T}$ ), we may take  $\lambda = 1$  from now on, and write  $p \equiv p_1 = \min_{0 \leq \gamma \leq 1} \bar{\rho}(\gamma)$ .

(a) Suppose  $p < 1$ . We will use the fact that the random walk is positive recurrent if and only if the conductances have finite sum [19, Propn. 9-131]. If  $\bar{\rho}(\gamma) < 1$  for some  $\gamma \in [0, 1]$  then

$$\mathbb{E} \left[ \sum_{v \in \mathcal{T}} C_v^\gamma \right] \asymp \sum_{n \geq 0} \bar{\rho}(\gamma)^n < \infty,$$

so  $\sum_{v \in \mathcal{T}} C_v^\gamma < \infty$  a.s. In particular  $C_v \leq C_v^\gamma < 1$  for all but finitely many  $v \in \mathcal{T}$  so  $\sum_{v \in \mathcal{T}} C_v < \infty$  a.s.

(b) Suppose  $p > 1$ . We will show that on the event of non-extinction there exists  $w < 1$  such that

$$\liminf_{n \rightarrow \infty} w^n \sum_{v \in D_n} C_v > 0; \quad (5.3)$$

transience then follows from [24, Cor. 4.2]. By the proof on [25, p. 129],

$$p = \max_{0 < y \leq 1} \left\{ y \inf_{\gamma \geq 0} y^{-\gamma} \bar{\rho}(\gamma) \right\},$$

and we fix  $y \in (0, 1]$  achieving this maximum. Then Lem. 5.2 implies that

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \bar{\mathcal{Q}}_{n\star}^a(C_{v_n} > z^n) \geq -\log[y\bar{\rho}(0)/p] \quad \forall z < y, \forall a \in \mathcal{Q}.$$

Therefore we can choose  $z < y$ ,  $\ell \in \mathbb{N}$ , and  $\epsilon, w \in (0, 1)$  such that

$$\min_{a \in \mathcal{Q}} \bar{\mathcal{Q}}_{\ell\star}^a(\{C_{v_\ell} > z^\ell\} \cap \{\alpha_w \geq \epsilon \forall o < w \leq v_\ell\}) \geq q > (wz\bar{\rho}(0))^{-\ell}.$$

Now consider the following percolation process (same as on [25, p. 130]): let  $\mathcal{T} \sim \text{MGW}(\cdot | \mathbb{X}^c)$ , and let  $\mathcal{T}[\ell]$  be the tree with vertices  $\{v \in \mathcal{T} : |v| \equiv 0 \pmod{\ell}\}$ , with an edge  $v \rightarrow w$  if and only if  $|w| = |v| + \ell$  in  $\mathcal{T}$ . Form a random subgraph  $\mathcal{T}[\ell]^{\text{perc}} \subseteq \mathcal{T}[\ell]$  by keeping the edge  $v \rightarrow w$  if and only if

$$\prod_{v < u \leq w} \alpha_u > z^\ell \quad \text{and} \quad \min_{v < u \leq w} \alpha_u \geq \epsilon,$$

in which case we write  $v \rightsquigarrow w$ . The subtree of  $\mathcal{T}[\ell]^{\text{perc}}$  descended from any vertex  $v$  has the law of a multi-type Galton–Watson tree with mean offspring numbers

$$A^{\text{perc}}(a, b) = \mathbb{E}_{\text{MGW}^a} \left[ \sum_{v \in D_\ell} \mathbf{1}_{\{\chi_v = b\}} \mathbf{1}_{\{o \rightsquigarrow v\}} \right] = \frac{e_a \bar{\rho}(0)^\ell}{e_b} \mathbb{E}_{\bar{\mathcal{Q}}_{\ell\star}^a} [\mathbf{1}_{\{\chi_{v_\ell} = b\}} \mathbf{1}_{\{o \rightsquigarrow v_\ell\}}].$$

We calculate

$$\sum_b A^{\text{perc}}(a, b) e_b \geq e_a \bar{\rho}(0)^\ell q > e_a (wz)^{-\ell},$$

so  $A^{\text{perc}}$  has Perron–Frobenius eigenvalue larger than  $(wz)^{-\ell}$ , and consequently  $\mathcal{T}[\ell]^{\text{perc}}$  a.s. has a connected component which is an infinite tree  $\mathcal{T}[\ell]^*$  of branching number larger than  $(wz)^{-\ell}$ , rooted at some  $o^* \in \mathcal{T}[\ell]^{\text{perc}}$ . It follows that the left-hand side of (5.3) is

$$\geq C_{o^*} w^{|o^*|^\ell} \epsilon^{\ell-1} \liminf_{n \rightarrow \infty} \sum_{v \in D_n(\mathcal{T}[\ell]^*)} (wz)^{n\ell} > 0$$

which concludes the proof.  $\square$

## 6. APPENDIX: GENERAL PROPERTIES OF MGW TREES

In this section we prove some basic facts about MGW trees which were used in the proof of the main theorem. In §6.2 we prove Propn. 3.1 which states that (H3<sup>p</sup>) implies  $\mathbb{E}_{\text{MGW}}[W_o^p] < \infty$ . In §6.3 we prove a conductance lower bound (Propn. 6.5) which gives Lem. 4.6. We begin in §6.1 by collecting some preliminary observations.

**6.1. Generating function and subtree of infinite descent.** Let

$$F(\underline{s}) \equiv (F^a(\underline{s}))_{a \in \mathcal{Q}} \equiv \left( \mathbb{E}_{\mathbf{q}^a} \left[ \prod_{b \in \mathcal{Q}} s_b^{x_b} \right] \right)_{a \in \mathcal{Q}}, \quad \underline{s} \in [0, 1]^{\mathcal{Q}};$$

we refer to  $F$  as the generating function of the MGW tree. If  $F^{(n)}$  denotes the  $n$ -fold composition of  $F$ , then for all  $a \in \mathcal{Q}$

$$\mathbb{E}_{\text{MGW}^a} \left[ \prod_{b \in \mathcal{Q}} s_b^{Z_n(b)} \right] = (F^{(n)}(\underline{s}))_a, \quad \text{MGW}^a(|\underline{Z}_n| = 0) = (F^{(n)}(\underline{0}))_a.$$

Next let

$$\Phi(s) \equiv (\Phi^a(s))_{a \in \mathcal{Q}} \equiv \left( \mathbb{E}_{\text{MGW}^a} [e^{-sW_o}] \right)_{a \in \mathcal{Q}}, \quad s \geq 0,$$

and let  $\phi(s) \equiv \langle \underline{g}, \Phi(s) \rangle = \mathbb{E}_{\text{MGW}}[e^{-sW_o}]$ . Since

$$\Phi(s) = \lim_{n \rightarrow \infty} F^{(n)}((\exp\{-se_b/\rho^n\})_{b \in \mathcal{Q}}),$$

we have the functional relation  $\Phi(s) = F[\Phi(s/\rho)]$ .

For many purposes the case of  $\text{MGW}(\mathbb{X}) \in (0, 1)$  can be reduced to the simpler case of an a.s. infinite tree without leaves by the following transformation which is discussed in [1, §I.12] for the single-type case. For  $\mathcal{T} \sim \text{MGW}$ , consider the subtree  $\mathcal{T}^\infty$  consisting of those vertices  $v$  of *infinite descent*, i.e. with  $|\mathcal{T}^{(v)}| = \infty$ . Conditioned on  $\mathbb{X}^c$ ,  $\mathcal{T}^\infty$  is an a.s. infinite tree without leaves, following a transformation of the original



MGW given by generating function  $\acute{F}(\underline{s}) = (\acute{F}^a(\underline{s}))_{a \in \mathcal{Q}}$ , where, with  $\mathfrak{r}_a \equiv \text{MGW}^a(\mathbb{X})$ ,

$$\begin{aligned} \acute{F}^a(\underline{s}) &\equiv \frac{1}{1 - \mathfrak{r}_a} \sum_{\underline{x}} \mathfrak{q}^a(\underline{x}) \prod_{b \in \mathcal{Q}} \sum_{y_b \leq x_b} \binom{x_b}{y_b} \mathfrak{r}_b^{x_b - y_b} (1 - \mathfrak{r}_b)^{y_b} s_b^{y_b} \\ &= \frac{F_a((\mathfrak{r}_b + (1 - \mathfrak{r}_b)s_b)_{b \in \mathcal{Q}})}{1 - \mathfrak{r}_a}. \end{aligned}$$

The transformed law has mean matrix

$$\acute{A} = D^{-1}AD, \quad D = \text{diag}((1 - \mathfrak{r}_a)_{a \in \mathcal{Q}}),$$

so in particular it has the same Perron–Frobenius eigenvalue as  $A$ . Finally, it is clear that if the original law satisfies (H3<sup>p</sup>) then so does the transformed law.

**6.2. Positive moments of the normalized population size.** In this section we show that moment conditions on the MGW offspring distribution translate directly to moment conditions on the normalized population size of the entire tree. We begin by recalling an easy fact concerning Laplace transforms (see e.g. [13, §XIII]).

**Lemma 6.1.** *Let  $\varphi(s) \equiv \mathbb{E}[e^{-sW}]$  be the Laplace transform of a non-negative random variable  $W$ . For any integer  $n \geq 0$ ,  $\mathbb{E}[W^n] < \infty$  if and only if there exist finite coefficients  $m_0, \dots, m_n$  such that*

$$\sum_{r=0}^n \frac{m_r}{r!} s^r = \varphi(s) + o(s^n), \quad s \downarrow 0. \quad (6.1)$$

In this case  $m_r = \mathbb{E}[(-W)^r] = \lim_{s \downarrow 0} \varphi^{(r)}(s)$ , and the left-hand side of (6.1) is the  $n$ -th order (one-sided) Taylor expansion  $P_{n,0}\varphi$  of  $\varphi$  at 0.

*Proof.* ( $\Rightarrow$ ) The function  $\varphi$  is infinitely differentiable on  $(0, \infty)$  with  $n$ -th derivative given by  $(-1)^n \varphi^{(n)}(s) = \mathbb{E}[e^{-sW} W^n]$ , and by the monotone convergence theorem

$$\varphi^{(n)}(0) \equiv \lim_{s \downarrow 0} \varphi^{(n)}(s) = \mathbb{E}[W^n] \in [0, \infty].$$

Writing  $\mathbf{e}(x) \equiv e^{-x}$ , by Taylor's theorem

$$R_{n,0}\mathbf{e}(x) \equiv (-1)^{n+1}[e^{-x} - P_{n,0}\mathbf{e}(x)] = \frac{x^n}{n!}(1 - e^{-\zeta}), \quad 0 \leq \zeta \leq x.$$

If  $\mathbb{E}[W^n] < \infty$  then

$$\mathbb{E}[R_{n,0}\mathbf{e}(sW)] = \frac{s^n}{n!} \mathbb{E}[W^n(1 - e^{-\zeta W})], \quad 0 \leq \zeta \leq s.$$

The right-hand side is  $o(s^n)$  so (6.1) holds with

$$\varphi(s) - \sum_{r=0}^n \frac{m_r}{r!} s^r = \varphi(s) - P_{n,0}\varphi(s) = (-1)^{n+1} \mathbb{E}[R_{n,0}\mathbf{e}(sW)]. \quad (6.2)$$

( $\Leftarrow$ ) Assuming (6.1), suppose inductively that  $m_r = \varphi^{(r)}(0)$  for  $0 \leq r \leq k$  with  $k < n$ . For  $0 < s_0 \leq s$  we have

$$\varphi(s) = P_{k,s_0}\varphi(s) + \frac{\varphi^{(k+1)}(\zeta)}{(k+1)!}(s-s_0)^{k+1}, \quad s_0 \leq \zeta \leq s,$$

so by (6.1) and the inductive hypothesis

$$o(s^{k+1}) = \varphi(s) - \sum_{r=0}^{k+1} \frac{m_r}{r!} s^r = \sum_{r=0}^k \frac{o(s_0)}{r!} + \frac{\varphi^{(k+1)}(\zeta)(s-s_0)^{k+1} - m_{k+1}s^{k+1}}{(k+1)!}.$$

Taking  $s_0 \ll s^{k+1}$  we find a contradiction unless  $\lim_{\zeta \downarrow 0} \varphi^{(k+1)}(\zeta) = m_{k+1}$ .  $\square$

**Lemma 6.2.** *If (H1), (H2) and (H3<sup>n</sup>) hold with  $n \in \mathbb{Z}_{\geq 2}$ , then  $\mathbb{E}_{\text{MGW}}[W_o^n] < \infty$ .*

*Proof.* Following the proof of [4, Thm. 0], we will show  $\mathbb{E}_{\text{MGW}}[W_o^n] < \infty$  using the characterization Lem. 6.1 of the derivatives at zero of the Laplace transform  $\phi(s) = \mathbb{E}_{\text{MGW}}[e^{-sW_o}]$ . Write  $\mathbb{S}_Q \equiv \{\underline{v} \in [0, \infty)^Q : \sum_{a \in Q} v_a = 1\}$ , and define

$$f(t; \underline{v}) \equiv \langle \underline{g}, F(e^{-t\underline{v}}) \rangle = \mathbb{E}_{\text{MGW}}[e^{-t\langle \underline{v}, \underline{Z}_1 \rangle}], \quad t \geq 0, \underline{v} \in \mathbb{S}_Q.$$

By (6.2),

$$f(t; \underline{v}) = P_{n,0}f(t; \underline{v}) + (-1)^{n+1}R_{n,0}f(t; \underline{v}), \quad \lim_{t \downarrow 0} \left( t^{-n} \sup_{\underline{v} \in \mathbb{S}_Q} R_{n,0}f(t; \underline{v}) \right) = 0,$$

where  $P_{n,0}f(t; \underline{v})$  is a polynomial of degree at most  $n$  in the entries of  $t\underline{v}$  satisfying

$$P_{n,0}f(t; \underline{v}) = 1 - t\langle \underline{v}, \mathbb{E}_{\text{MGW}}[\underline{Z}_1] \rangle + O(t^2) = 1 - \rho t\langle \underline{v}, \underline{g} \rangle + O(t^2). \quad (6.3)$$

If we let  $t \equiv t(s) \geq 0$  and  $\underline{v} \equiv \underline{v}(s) \in \mathbb{S}_Q$  be defined by  $\Phi(s/\rho) = e^{-t\underline{v}}$ , then

$$\phi(s) = \langle \underline{g}, \Phi(s) \rangle = \langle \underline{g}, F[\Phi(s/\rho)] \rangle = f(t; \underline{v}) = P_{n,0}f(t; \underline{v}) + o(t^n) \quad (6.4)$$

(using that  $R_{n,0}f(t; \underline{v}) = o(t^n)$  uniformly over  $\underline{v} \in \mathbb{S}_Q$ ).

We next expand  $t\underline{v}$  in powers of  $s$ . Note that  $\phi(s) - m_0 - m_1s = o(s)$  where  $m_0 = 1$ ,  $m_1 = -\mathbb{E}_{\text{MGW}}[W_o]$ , so suppose inductively that  $\mathbb{E}_{\text{MGW}}[W_o^k] < \infty$  for some  $1 < k < n$ . Lem. 6.1 implies the existence of polynomials  $q_k^a(s) = e_a + O(s)$  such that

$$tv_a = -\log \Phi^a(s/\rho) = s q_k^a(s) + o(s^k), \quad s \downarrow 0. \quad (6.5)$$

By recalling (6.4) and comparing (6.3) against the Taylor expansion of  $\rho\langle \underline{g}, e^{-t\underline{v}} \rangle$ , we find

$$[\phi(s) - 1] - \rho[\phi(s/\rho) - 1] = Q_n(t\underline{v}) + o(t^n), \quad t \downarrow 0.$$

for  $Q_n : \mathbb{R}^Q \rightarrow \mathbb{R}$  polynomial with  $Q_n(t\underline{v}) = O(t^2)$ . But squaring (6.5) gives  $(tv_a)^2 = s^2 q_k^a(s)^2 + o(s^{k+1})$ , and substituting into the above and dividing through by  $s$  gives that

$$\psi(s) - \psi(s/\rho) \equiv \frac{\phi(s) - 1}{s} - \frac{\phi(s/\rho) - 1}{s/\rho} = sP_k(s) + o(s^k)$$

for  $P_k$  polynomial in  $s$ . Since  $\rho > 1$  and  $\lim_{s \downarrow 0} \psi(s) = \phi'(0)$ ,

$$\psi(s) = \psi'(0) + \sum_{j \geq 0} (s/\rho^j) P_n(s/\rho) + o(s^k) = \tilde{P}_n(s) + o(s^k)$$

for  $\tilde{P}_n$  another polynomial in  $s$ . This verifies the inductive hypothesis by the definition of  $\psi$  together with another application of Lem. 6.1.  $\square$

If  $(H3^n)$  holds with  $n \in \mathbb{Z}_{\geq 2}$  then we write  $\phi(s) = P_{n,0}\phi(s) + (-1)^{n+1}R_{n,0}\phi(s)$  with  $P_{n,0}\phi$  polynomial of degree at most  $n$  and  $R_{n,0}\phi(s) = o(s^n)$ . An easy consequence of the proof of the Lem. 6.2 is the following

**Corollary 6.3.** *If  $(H3^n)$  holds with  $n \in \mathbb{Z}_{\geq 2}$  then*

$$R_{n,0}\phi(s) - \rho R_{n,0}\phi(s/\rho) = R_{n,0}f(t, \underline{v}) + O(t^{n+1})$$

for  $\Phi(s/\rho) \equiv e^{-t\underline{v}}$ .

*Proof.* Summing (6.5) over  $a \in \mathcal{Q}$  gives the existence of  $b_2, \dots, b_n$  finite such that

$$t = s + \sum_{r=2}^n b_r s^r + o(s^n).$$

It follows easily that  $s$  has a similar expansion in terms of  $t$ : indeed  $s = t + o(t)$ , so suppose inductively that for some  $1 \leq k < n$  there exist  $c_2, \dots, c_k$  finite such that  $s = t + \sum_{r=2}^k c_r t^r + o(t^k)$ . Then

$$s = t - \sum_{r=2}^n b_r \left( t + \sum_{r=2}^k c_r t^r + o(t^k) \right)^2 + o(s^n),$$

which is a polynomial in  $t$  plus  $o(t^{k+1})$ . This verifies the inductive hypothesis so we conclude that  $s = t + \sum_{r=2}^n c_r t^r + o(t^n)$  as claimed. From the proof of Lem. 6.2 we have

$$\begin{aligned} o(s^n) &= (-1)^{n+1} [R_{n,0}\phi(s) - \rho R_{n,0}\phi(s/\rho)] \\ &= f(t; \underline{v}) - \rho \langle \underline{g}, e^{-t\underline{v}} \rangle - [P_{n,0}\phi(s) - \rho P_{n,0}\phi(s/\rho)] \\ &= f(t; \underline{v}) - Q_n^\backslash(t\underline{v}) - s^2 q(s) + O(t^{n+1}) \end{aligned}$$

for  $Q_n^\backslash$  and  $q$  polynomial. But by the above  $s^2$  can be expressed as a polynomial in  $t$  up to  $o(t^{n+1})$  error, so in fact

$$o(t^n) = (-1)^{n+1} [R_{n,0}\phi(s) - \rho R_{n,0}\phi(s/\rho)] = f(t; \underline{v}) - Q_n^\backslash(t\underline{v}) + O(t^{n+1})$$

for  $Q_n^\backslash : \mathbb{R}^{\mathcal{Q}} \rightarrow \mathbb{R}$  polynomial in  $t\underline{v}$  of degree at most  $n$  in  $t$ , whence necessarily  $Q_n^\backslash(t\underline{v}) = P_{n,0}f(t; \underline{v})$  as claimed.  $\square$

If  $\Phi(s/\rho) = e^{-t\underline{v}}$  with  $t \equiv t(s) \geq 0$ ,  $\underline{v} \equiv \underline{v}(s) \in \mathbb{S}_{\mathcal{Q}}$  as above, then

$$t = - \sum_{a \in \mathcal{Q}} \log \mathbb{E}_{\text{MGW}^a} [e^{-s W_o}].$$

In particular  $t'(s)$  is finite and positive for all  $s > 0$  with  $\lim_{s \downarrow 0} t'(s) = 1$ , so  $s \mapsto t(s)$  is an increasing bijection from  $[0, \infty)$  to  $[0, t_{\max})$  where  $t_{\max} = - \sum_{a \in \mathcal{Q}} \log \text{MGW}^a(\mathbb{X})$ . For  $t < t_{\max}$  we therefore write  $\underline{v}_t \equiv \underline{v}(s)$  with  $s$  defined by  $t = t(s)$ .

*Proof of Propn. 3.1.* Since the subtree of infinite descent described in §6.1 has the same normalized population size as the original tree, we may reduce to the case  $\text{MGW}(\mathbb{X}) = 0$  so that  $t_{\max} = \infty$ .

By Lem. 6.2 we may take  $p = n + \beta$  for  $\beta \in (0, 1)$ , and by [4, Thm. B] the result follows upon showing

$$\int_0^1 R_{n,0} \phi(s) s^{-(1+p)} ds < \infty.$$

By the proof of [4, Propn. 5], this in turn follows upon showing

$$\int_0^1 R_{n,0} f(t; \underline{v}_t) t^{-(1+p)} dt < \infty. \quad (6.6)$$

(replacing [4, (3.13)] with (6.6) and [4, (3.9)] with Cor. 6.3). Recalling (6.2),

$$\int_0^1 \frac{R_{n,0} f(t; \underline{v}_t)}{t^p} \frac{dt}{t} \leq \frac{1}{n!} \int_0^1 \frac{\mathbb{E}_{\text{MGW}}[|\underline{Z}_1|^n (1 - e^{-t|\underline{Z}_1|})]}{t^\beta} \frac{dt}{t}.$$

Applying Fubini's theorem and making the change of variable  $t \mapsto |\underline{Z}_1|t$  gives that the above is

$$= \frac{1}{n!} \mathbb{E}_{\text{MGW}} \left[ |\underline{Z}_1|^p \int_0^{|\underline{Z}_1|} \frac{1 - e^{-t}}{t^\beta} \frac{dt}{t} \right] < \infty,$$

which concludes the proof.  $\square$

**6.3. Harmonic moments and conductance estimates.** In this section we prove the existence of harmonic moments for the normalized population size, extending part of [29, Thm. 1] to the multi-type setting (using a similar proof). Using this result we adapt the methods of [30, Lem. 2.2] to prove the conductance estimates used in the proofs of Cor. 4.7 and Lem. 4.8.

**Lemma 6.4.** *Assume (H1) and (H2). There exists some  $r > 0$  for which*

$$\mathbb{E}_{\text{MGW}}[W_o^{-r} | \mathbb{X}^c] \leq \limsup_{n \rightarrow \infty} \mathbb{E}_{\text{MGW}}[\mathfrak{Z}_n^{-r} | \mathbb{X}^c] < \infty.$$

*Proof.* Since the subtree of infinite descent described in §6.1 has the same normalized population size as the original tree, we may reduce to the case  $\text{MGW}(\mathbb{X}) = 0$ . Expanding  $F^{(n)}(\underline{s})$  as a power series in  $\underline{s}$  we find

$$(F^{(n)}(\underline{s}))_a \leq \text{MGW}^a(|\underline{Z}_n| = 1) \|\underline{s}\|_\infty + \text{MGW}^a(|\underline{Z}_n| > 1) \|\underline{s}\|_\infty^2.$$

By (H1), there exists  $n_0$  such that  $\min_{a \in \Omega} \text{MGW}^a(|\underline{Z}_n| > 1) > 0$  for all  $n \geq n_0$ , so that  $F^{(n_0)}$  is a contraction on  $[0, s_0]^Q$  for any  $s_0 < 1$ . By iterating this estimate, for any  $s_0 < 1$  there exist constants  $C < \infty$  and  $\gamma < 1$  such that

$$\|F^{(n)}(\underline{s})\|_\infty \leq C\gamma^n \|\underline{s}\|_\infty.$$

By Fubini's theorem,

$$\mathbb{E}_{\text{MGW}}[\mathfrak{Z}_n^{-r}] = \frac{\rho^{nr}}{\Gamma(r)} \int_0^\infty f_n(u) u^{r-1} du, \quad f_n(u) \equiv \mathbb{E}_{\text{MGW}}[e^{-u\langle \underline{Z}_n, \underline{e} \rangle}].$$

We break up the integral into three parts, writing  $I_{a,b}$  for the integral over  $[a, b]$ : By a change of variables,

$$\Gamma(r) I_{0,\rho^{-n}} = \int_0^1 \mathbb{E}_{\text{MGW}}[e^{-u\mathfrak{Z}_n}] u^{r-1} du \leq \frac{1}{r}.$$

Next, we have

$$f_n(u) = \langle \underline{g}, F^{(n)}(e^{-ue_b}) \rangle \leq \|F^{(n)}(e^{-ue_b})\|_\infty \leq C\gamma^n \|(e^{-ue_b})_{b \in \Omega}\|_\infty \leq C\gamma^n e^{-ue_{\min}},$$

where for any  $u_0 > 0$  we may choose constants  $C < \infty$  and  $\gamma < 1$  uniformly over all  $u \geq u_0$ . Therefore

$$\Gamma(r) I_{1,\infty} \leq C(\rho^r \gamma)^n \int_1^\infty e^{-ue_{\min}} u^{r-1} du.$$

For  $r > 0$  small enough so that  $\gamma\rho^r < 1$ , we have  $\lim_{n \rightarrow \infty} I_{1,\infty} = 0$ . It remains to consider

$$\begin{aligned} \Gamma(r) I_{\rho^{-n},1} &= \rho^{nr} \sum_{i=1}^n \int_{1/\rho^i}^{1/\rho^{i-1}} f_n(u) u^{r-1} du \\ &= \sum_{i=1}^n \rho^{r(n-i)} \int_1^\rho \mathbb{E}_{\text{MGW}}[e^{-u\langle Z_n, \mathbf{e} \rangle / \rho^i}] u^{r-1} du. \end{aligned}$$

By conditioning on the first  $n-i$  levels of the tree,

$$\begin{aligned} \mathbb{E}_{\text{MGW}}[e^{-u\langle Z_n, \mathbf{e} \rangle / \rho^i}] &= \mathbb{E}_{\text{MGW}} \left[ \prod_{v \in D_{n-i}} \mathbb{E}_{\text{MGW}}[e^{-u\langle Z_v^v, \mathbf{e} \rangle / \rho^i} \mid \chi_v] \right] \\ &= \mathbb{E}_{\text{MGW}} \left[ \prod_{a \in \Omega} \Phi_i^a(u)^{Z_{n-i}(a)} \right] = \langle \underline{g}, F^{(n-i)}[(\Phi_i^a(u))_{a \in \Omega}] \rangle, \end{aligned}$$

where  $\Phi_i^a(u) \equiv \mathbb{E}_{\text{MGW}^a}[e^{-u\mathfrak{Z}_i}]$ . But

$$\sup_{i \geq 1} \sup_{u \geq 1} \Phi_i^a(u) = \sup_{i \geq 1} \Phi_i^a(1) < 1,$$

since  $\Phi_i^a(1) < 1$  for all  $a, i$  and  $\Phi_i^a(1) \rightarrow \mathbb{E}_{\text{MGW}^a}[e^{-uW_o}]$  which is less than 1 by the Kesten–Stigum theorem as noted in §2.3 (using (H2)). Therefore

$$\Gamma(r) I_{\rho^{-n},1} \leq C \sum_{i=1}^n (\gamma\rho^r)^{n-i} \int_1^\rho u^{r-1} du,$$

which is bounded in  $n$  for small enough  $r$ . Putting the estimates together concludes the proof.  $\square$

We conclude with the following conductance lower bound, a version of [30, Lem. 2.2]. This clearly implies Lem. 4.6 which was used in the proof of the annealed MGW-CLT Propn. 4.3.

**Proposition 6.5.** *(a) Under (H1), (H2), and (H3<sup>2</sup>), there exist  $0 < r, C < \infty$  such that for all  $\epsilon > 0$ ,  $\text{MGW}(\mathcal{C}_{o,k}^{-1} \geq k^{1+\epsilon} \mid \mathbb{X}^c) \leq Ck^{-r\epsilon}$ .*

(b) If further (H3<sup>p</sup>) holds with  $p > 2$ , then for MGW-a.e.  $\mathcal{T} \notin \mathbb{X}$  there exists a random constant  $C_{\mathcal{T}} < \infty$  such that  $\mathcal{C}_{o,k}^{-1} \leq C_{\mathcal{T}}k$  for all  $k$ .

*Proof.* (a) Recall that a *unit flow* is a non-negative function  $U$  on the vertices of  $\mathcal{T}$  such that for all  $v \in \mathcal{T}$ ,  $U(v) = \sum_{w \in \partial^+ v} U(w)$ . For  $v \in D_\ell$  define

$$U(v) = \frac{W_v}{\sum_{u \in D_\ell} W_u} = \frac{W_v}{\rho^\ell W_o};$$

it is easily seen that  $U$  is a well-defined unit flow on  $\mathbb{X}^c$ . It gives positive flow only to vertices of infinite descent, so by the discussion of §6.1 we may reduce to the case  $\text{MGW}(\mathbb{X}) = 0$ . By Thomson's principle [28, §2.4]

$$\mathcal{C}_{o,k}^{-1} \leq \sum_{\ell=1}^k \rho^\ell \sum_{v \in D_\ell} U(v)^2 = \frac{1}{W_o^2} \sum_{\ell=1}^k \frac{1}{\rho^\ell} \sum_{v \in D_\ell} W_v^2.$$

By Hölder's inequality,

$$\mathbb{E}_{\text{MGW}}[\mathcal{C}_{o,k}^{-r}] \leq \mathbb{E}_{\text{MGW}} \left[ \sum_{\ell=1}^k \frac{1}{\rho^\ell} \sum_{v \in D_\ell} W_v^2 \right]^r \mathbb{E}[W_o^{-2r/(1-r)}]^{1-r} \leq Ck^r$$

for  $r$  sufficiently small, using Lem. 6.4 and  $p \geq 2$ . It follows from Markov's inequality that  $\text{MGW}(\mathcal{C}_{o,k}^{-1} \geq k^{1+\epsilon}) \leq Ck^{-r\epsilon}$ .

(b) We claim there exist  $0 < c, c' < \infty$  deterministic such that

$$\text{MGW} \left( \frac{1}{|D_k|} \sum_{v \in D_k} W_v^2 \geq c' \right) \leq \rho^{-ck} \quad (6.7)$$

Assuming the claim, we have

$$\rho^k \sum_{v \in D_k} U(v)^2 = \frac{1}{\rho^k} \frac{\sum_{v \in D_k} W_v^2}{W_o^2} \leq \frac{C3_k}{W_o^2} \left( \frac{1}{|D_k|} \sum_{v \in D_k} W_v^2 \right),$$

so by Borel–Cantelli

$$\limsup_{k \rightarrow \infty} \rho^k \sum_{v \in D_k} U(v)^2 \leq \frac{Cc'}{W_o} < \infty,$$

which by Thomson's principle implies  $\mathcal{C}_{o,k}^{-1} \leq C_{\mathcal{T}}k$ .

It remains to prove (6.7). For any  $1 \leq 1+r \leq 2 \wedge (p/2)$ , Lem. 3.7 and Markov's inequality give

$$\begin{aligned} & \text{MGW} \left( \left| \frac{1}{|D_k|} \sum_{v \in D_k} (W_v^2 - \mathbb{E}_{\text{MGW}^{x_v}}[W_v^2]) \right| \geq \epsilon \mid \mathcal{F}_k \right) \\ & \leq \frac{2}{(\epsilon|D_k|)^{1+r}} \sum_{v \in D_k} \mathbb{E}_{\text{MGW}^{x_v}}[|W_v^2 - \mathbb{E}_{\text{MGW}^{x_v}}[W_v^2]|^{1+r}] \leq C(\epsilon, r)|D_k|^{-r}. \end{aligned}$$

Taking expectations and applying Lem. 6.4 then gives

$$\text{MGW} \left( \left| \frac{1}{|D_k|} \sum_{v \in D_k} W_v^2 - \frac{1}{|D_k|} \sum_{v \in D_k} \mathbb{E}_{\text{MGW}^{\chi_v}} [W_v^2] \right| \geq \epsilon \right) \leq C(\epsilon, r) \rho^{-rk}$$

for  $r$  sufficiently small. But

$$\frac{1}{|D_k|} \sum_{v \in D_k} \mathbb{E}_{\text{MGW}^{\chi_v}} [W_v^2]$$

is clearly bounded uniformly in  $k$  by a deterministic constant, so (6.7) is proved.  $\square$

## REFERENCES

- [1] K. B. Athreya and P. E. Ney. *Branching processes*. Dover Publications Inc., Mineola, NY, 2004. Reprint of the 1972 original [Springer, New York; MR0373040].
- [2] G. Ben Arous, A. Fribergh, N. Gantert, and A. Hammond. Biased random walks on Galton–Watson trees with leaves. *Ann. Probab.*, 40(1), 2012.
- [3] P. Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, 1995. A Wiley-Interscience Publication.
- [4] N. H. Bingham and R. A. Doney. Asymptotic properties of supercritical branching processes. I. The Galton–Watson process. *Advances in Appl. Probability*, 6:711–731, 1974.
- [5] E. Bolthausen and A.-S. Sznitman. On the static and dynamic points of view for certain random walks in random environment. *Methods Appl. Anal.*, 9(3):345–375, 2002. Special issue dedicated to Daniel W. Stroock and Srinivasa S. R. Varadhan on the occasion of their 60th birthday.
- [6] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004.
- [7] M. Bramson, P. Ney, and J. Tao. The population composition of a multitype branching random walk. *Ann. Appl. Probab.*, 2(3):575–596, 1992.
- [8] D. A. Croydon, A. Fribergh, and T. Kumagai. Biased random walk on critical Galton–Watson trees conditioned to survive. Preprint, [arXiv:1203.4078](https://arxiv.org/abs/1203.4078), 2012.
- [9] A. Dembo and O. Zeitouni. *Large deviations: techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998.
- [10] R. Durrett. *Probability: theory and examples*. Duxbury Press, Belmont, CA, 3rd edition, 2004.
- [11] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.
- [12] G. Faraud. A central limit theorem for random walk in a random environment on marked Galton–Watson trees. *Electron. J. Probab.*, 16(6):174–215, 2011.
- [13] W. Feller. *An introduction to probability theory and its applications. Vol. II*. John Wiley & Sons Inc., New York, 2nd edition, 1971.
- [14] T. E. Harris. *The theory of branching processes*. Dover Phoenix Editions. Dover Publications Inc., Mineola, NY, 2002. Corrected reprint of the 1963 original [Springer, Berlin; MR0163361 (29 #664)].
- [15] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.
- [16] Y. Hu and Z. Shi. A subdiffusive behaviour of recurrent random walk in random environment on a regular tree. *Probab. Theory Related Fields*, 138(3-4):521–549, 2007.
- [17] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.
- [18] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.



- [19] J. G. Kemeny, J. L. Snell, and A. W. Knapp. *Denumerable Markov chains*. Springer-Verlag, New York, second edition, 1976.
- [20] H. Kesten and B. P. Stigum. A limit theorem for multidimensional Galton–Watson processes. *Ann. Math. Statist.*, 37:1211–1223, 1966.
- [21] T. Kurtz, R. Lyons, R. Pemantle, and Y. Peres. A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes. In *Classical and modern branching processes (Minneapolis, MN, 1994)*, volume 84 of *IMA Vol. Math. Appl.*, pages 181–185. Springer, New York, 1997.
- [22] A. E. Kyprianou and A. Rahimzadeh Sani. Martingale convergence and the functional equation in the multi-type branching random walk. *Bernoulli*, 7(4):593–604, 2001.
- [23] Q. Liu. On generalized multiplicative cascades. *Stochastic Process. Appl.*, 86(2):263–286, 2000.
- [24] R. Lyons. Random walks and percolation on trees. *Ann. Probab.*, 18(3):931–958, 1990.
- [25] R. Lyons and R. Pemantle. Random walk in a random environment and first-passage percolation on trees. *Ann. Probab.*, 20(1):125–136, 1992.
- [26] R. Lyons, R. Pemantle, and Y. Peres. Ergodic theory on Galton–Watson trees: speed of random walk and dimension of harmonic measure. *Ergodic Theory Dynam. Systems*, 15(3):593–619, 1995.
- [27] R. Lyons, R. Pemantle, and Y. Peres. Biased random walks on Galton–Watson trees. *Probab. Theory Related Fields*, 106(2):249–264, 1996.
- [28] R. Lyons and Y. Peres. *Probability on Trees and Networks*. **Preprint**, 2009.
- [29] P. E. Ney and A. N. Vidyashankar. Harmonic moments and large deviation rates for supercritical branching processes. *Ann. Appl. Probab.*, 13(2):475–489, 2003.
- [30] R. Pemantle and Y. Peres. Galton–Watson trees with the same mean have the same polar sets. *Ann. Probab.*, 23(3):1102–1124, 1995.
- [31] Y. Peres and O. Zeitouni. A central limit theorem for biased random walks on Galton–Watson trees. *Probab. Theory Related Fields*, 140(3-4):595–629, 2008.
- [32] K. Petersen. *Ergodic theory*, volume 2 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1989. Corrected reprint of the 1983 original.
- [33] V. V. Petrov. *Sums of independent random variables*. Springer-Verlag, New York, 1975. Translated from the Russian by A. A. Brown, *Ergebnisse der Mathematik und ihrer Grenzgebiete*, Band 82.
- [34] D. Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.
- [35] O. Zeitouni. Random walks in random environment. In *Lectures on probability theory and statistics*, volume 1837 of *Lecture Notes in Math.*, pages 189–312. Springer, Berlin, 2004.

\*DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY  
BUILDING 380, SLOAN HALL, STANFORD, CALIFORNIA 94305

\*<sup>†</sup>DEPARTMENT OF STATISTICS, STANFORD UNIVERSITY  
SEQUOIA HALL, 390 SERRA MALL, STANFORD, CALIFORNIA 94305